JUL 5 1951

# The Journal of Experimental Education

A periodical report of scientific investigations relating to child development, curriculum, learning, teaching, supervision, measurements, statistics, and experimental techniques.

Volume XIX June, 1951 Number 4

CONTENTS Page

The Preferential Training Needs Record: A Study of In-Service Educational Needs of Teachers of the Atlanta Area Teacher Educational Service:
*Robert Travis Osborne* 271

Teacher Morale and Curriculum Development: A Statistical Analysis of Responses to a Reaction Inventory:
*William E. Coffman* 305

Practices Associated with Effective Discipline: A Descriptive Statistical Study of Discipline:
*Sidney L. Celler* 333

INDEX

$5.00 A YEAR PUBLISHED QUARTERLY $1.50 A COPY

Published by Dembar Publications, Inc.,
Madison 3, Wisconsin
Entered as second-class matter October 17, 1938 at the post office at Madison, Wisconsin, under the act of March 3, 1879.

# EDITORIAL BOARD


A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison 6, Wis.

Arthur T. Jersild, Professor of Education, Teachers College, Columbia University, New York City. Editorially responsible for materials on child welfare, guidance, and development, published each December.

Palmer O. Johnson, Professor of Education, University of Minnesota, Minneapolis, Minnesota. Editorially responsible for materials on measurements, statistics, and methods of experimental research, published each March.

H. H. Remmers, Professor of Educational Psychology, Director Division Educational Reference, Purdue University, Lafayette, Indiana. Editorially responsible for materials on learning, teaching and supervision, published each September.

J. Wayne Wrightstone, Director, Bureau of Educational Research, Board of Education of the City of New York, Brooklyn, New York, 110 Livingston Street, Brooklyn, New York. Editorially responsible for materials on curriculum construction, published each June.


# CONTRIBUTING EDITORS


Emmett A. Betts, Research Professor and Director of Reading Clinic, Pennsylvania State College, State College, Pennsylvania.

Gilbert L. Betts, Supervisor of Graduate Research in Education, Colorado State College, Fort Collins, Colorado.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Oscar K. Buros, Associate Professor of Education, Rutgers University, New Brunswick, New Jersey.

Guy T. Buswell, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Harold D. Carter, Associate Professor of Education, University of California, Berkeley 4, California.

Leslie L. Chisholm, Associate Professor of Education, State College of Washington, Pullman, Washington.

Herbert S. Conrad, Technical Consultant, College Entrance Examination Board, Princeton, New Jersey.

Stephen M. Corey, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Edward E. Cureton, 853 East Brewster Road, Alcoa, Tennessee.

Robert A. Davis, Professor of Education, Director of Bureau of Educational Research, University of Colorado, Boulder, Colorado.

Harl R. Douglass, Director of College of Education, University of Colorado, Boulder, Colorado.

Harold A. Edgerton, Director, Occupational Opportunities Service, Professor of Psychology, Ohio State University, Columbus 10, Ohio.

John C. Flanagan, Professor of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania.

Carter V. Good, Dean, Teachers College, University of Cincinnati, Cincinnati 21, Ohio.

Robert W. B. Jackson, Assistant Professor of Educational Research; Assistant Director, Department of Educational Research, Ontario College of Education, University of Toronto, Toronto, Canada.

Harold E. Jones, Professor of Psychology and Director, Institute of Child Welfare, University of California, Berkeley 4, California.

Noel Keys, Professor of Education and Lecturer in Human Relations, University of California, Berkeley, California.

D. Welty Lefever, Professor of Education, University of Southern California, Los Angeles, California.

Edward A. Lincoln, Consulting Psychologist, Halifax, Massachusetts.

Irving Lorge, Professor of Education, Executive Officer, Institute of Psychological Research, Teachers College, Columbia University, New York 27, New York.

A. R. Mead, Director of Educational Research, University of Florida, 330 P. K. Younge Building, Gainesville, Florida.

T. E. Newland, Lt. Comdr., USNR, 2702 Wisconsin Avenue, N. W., Washington 7, D. C.

C. W. Odell, Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Professor of Education, Director of Research in Child Development, University of Michigan, Ann Arbor, Michigan.

W. E. Peik, Dean and Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Walworth R. Plumb, Chairman, Division of Education and Psychology, University of Minnesota (Branch), Duluth, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

Clarence E. Ragsdale, Professor of Education, University of Wisconsin, Madison, Wisconsin.

William Reitz, Associate Professor of Education, College of Education Examiner, Wayne University, Detroit 2, Michigan.

Henry D. Rinsland, Professor of Education and Director of Educational Research, The University of Oklahoma, Norman, Oklahoma.

Robert T. Rock, Jr., Professor of Psychology, Head of Dept. of Psychology, Graduate School, Fordham University, New York City.

Phillip J. Rulon, Professor of Education, Harvard Graduate School of Education, Cambridge 38, Massachusetts.

Douglas E. Scates, Professor of Education, Duke University, Durham, North Carolina.

John Schmid, Board of Examiners, Michigan State College, East Lansing, Michigan.

Harold Seashore, Director, Test Division, The Psychological Corporation, New York 18, New York.

David Segel, Educational Consultant, Specialist in Tests and Measurements, Federal Security Agency, U. S. Office of Education, Washington, D. C.

Paul W. Terry, Professor of Educational Psychology, University of Alabama, University, Alabama.

Helen Thompson, Associate Attending Psychologist, New York Post-Graduate Hospital, 303 East 20th Street New York 3, N.Y.

Robert L. Thorndike, Associate Professor of Education, Teachers College, Columbia University, New York City.

Herbert A. Toops, Professor of Psychology, Ohio State University, Columbus, Ohio.

Maurice E. Troyer, Director, Bureau of School Services, Syracuse University, Syracuse 10, New York.

Helen M. Walker, Professor of Education, Teachers College, Columbia University, New York City.

Beth L. Wellman, Professor of Psychology, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

Guy M. Wilson, Emeritus Professor of Education Boston University, 33 Pine Street, Wellesley Hills, Massachusetts.

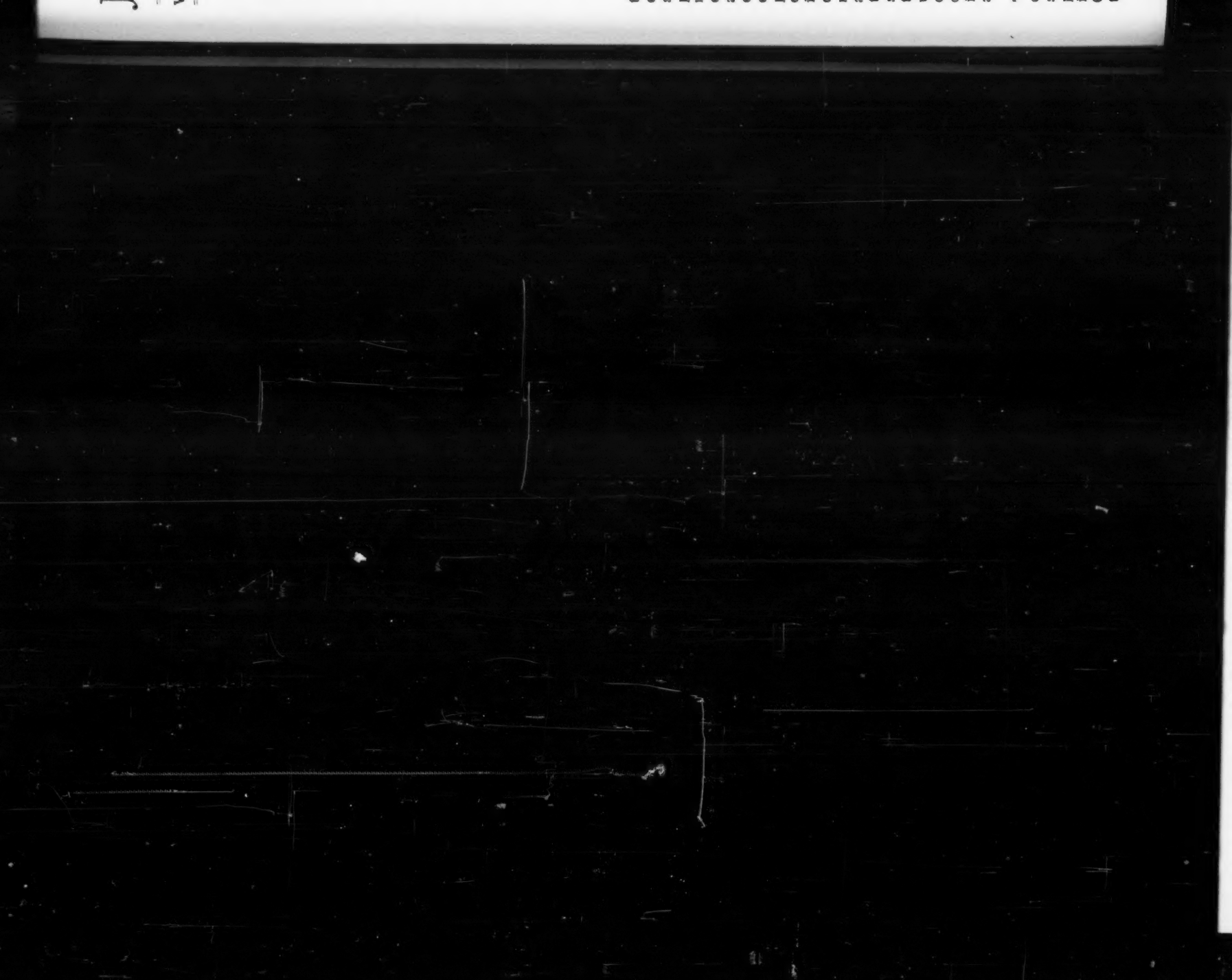
Paul A. Witty, Professor of Education, Director of Psycho-Educational Clinic, School of Education, Northwestern University, Evanston, Illinois.

Ernest R. Wood, Professor of Education, New York University, New York City.

D. A. Worcester, Chairman Department of Educational Psychology and Measurement, University of Nebraska, Lincoln, Nebraska.



LITHOGRAPHED BY LITHO PRODUCTIONS - MADISON, WISCONSIN



Journal of Experimental Education

Volume XIX June, 1951 Number 4

# THE PREFERENTIAL TRAINING NEEDS RECORD: A STUDY OF IN-SERVICE EDUCATIONAL NEEDS OF TEACHERS OF THE ATLANTA AREA TEACHER EDUCATION SERVICE

ROBERT TRAVIS OSBORNE
University of Georgia
Athens, Georgia

## SECTION I

## INTRODUCTION

Until recently, the conception of in-service teacher education has been limited to a formal course and credit program of teacher training similar to, if not identical with, the pre-service program, but offered for the express purpose of maintaining or raising levels of teacher certification. Such a program, of course, implied that after a teacher had taken a prescribed number of college courses or had acquired certain certificate requirements, her education would be complete. As a result of significant research in recent years, this conception of in-service training of teachers has been replaced by a program definitely slanted to meet the professional training needs felt by teachers. The teachers themselves are playing an increasingly important role in setting up and directing educational clinics, workshops, and other flexible cooperative in-service training activities. In the report "Teacher Education In Service" (23), the Commission on Teacher Education found this newer emphasis in teacher training to be on a continuous program of in-service training designed to lead to personal and professional growth.

Aided and encouraged by the Commission on Teacher Education, school authorities and teacher-training institutions are recognizing their responsibilities in this field and are attempting to provide experiences designed to result in self-improvement and in-service growth of teachers. C. E. Prall, in the American Council on Education publication, "State Programs for the Improvement of Teacher Education" (22), reports in detail the statewide cooperative progress of in-service training in Kentucky and Georgia. These two programs were described, as it was the feeling of the Commission that they were typical of others found throughout the country. The Kentucky program of in-service education for school and college personnel is a long-range, continuous process specifically aimed at improving education on-the-job. It involves educating staff members of teacher-training institutions as well as in-service training for teachers in the lower schools. The Georgia program was similar to other state programs but had for one of its main objectives "developing continuity in the educational guidance of teachers" (22). The following notes from the Committee report indicate, in part, how the Georgia program was implemented:

1. Plan profitable summer school experiences for teachers.
2. Encourage local principals and superintendents to assume responsibility for guiding the growth of classroom teachers.
3. Encourage study in local schools and community groups on local needs and ways of meeting them.
4. Cooperate with the state supervisors and local school teachers in planning and conducting educational clinics.
5. Set up workshops for principals and superintendents in cooperation with the high

school principals' association and the accrediting committees of the state.

6. Enlist the cooperation of organized professional groups in a common attack upon conditions affecting the economic status of teachers.

The newer trends in in-service training of teachers, as described for Kentucky and Georgia, are found also in numerous other states and communities.

An outstanding example of community cooperation on the problem of in-service teacher training is the Atlanta Area Teacher Education Service. In the Atlanta area, six institutions of higher learning, three county school systems, and three city school systems pooled their resources in order to provide in-service training activities with increasing attention being given to the personal, social, and professional needs of the teacher.

## SECTION II

## THE PROBLEM

The Atlanta Area Teacher Education Service was organized in March, 1945, as a result of recommendations to the Advisory Faculty Council of the University Center in Georgia by its Graduate Committee. Late in 1945, final approval of this recommendation was given by the Council of Presidents of the participating institutions of the University Center. At the time the Atlanta Area Teacher Education Service was organized, the institutions of higher learning participating in the University Center were: Agnes Scott College, Atlanta Art Association, Columbia Theological Seminary, Georgia Institute of Technology, Emory University and the University of Georgia. Participating public schools were: Atlanta, Decatur, and Marietta city systems, and the Cobb, DeKalb, and Fulton county systems.

"The purpose of the Atlanta Area Teacher Education Service is to combine the resources of the institutions with those of the public school systems in locating and attempting to solve the problems confronted by teachers and administrators of the public school systems in the area. Attempts to realize this purpose have been made through three specific types of activities. First, faculty members from the various institutions conduct classes which are taken by teachers for credit so that teaching certificates may be strengthened. Second, staff members of the various institutions serve as consultants to groups interested in securing help on a given problem in order that instruction may be improved. Third, specialists are brought in to the Atlanta area from other institutions to give special assistance in the field which they represent." (32)

The Atlanta Area Teacher Education Service intends to improve instruction in the elementary and secondary schools of the Atlanta area by helping teachers with their on-the-job problems. This program uses the general approach of giving teachers the kind of training experiences they want. It is believed that such an approach can be successful if the teachers know what training experiences they want, if they are conscious of their strengths and deficiencies, and if their administrators and the college instructors of the Atlanta Area Teacher Education Service are well-informed concerning the wants and the deficiencies of the personnel being served.

The purpose of this study is to determine what relationships exist between (1) the felt training needs of the teachers participating in the Atlanta Area Teacher Education Service, (2) the training experiences which the college instructors think the teachers need, and (3) the training experiences the teachers' administrators think the teachers need.

An earlier study (24) by Louis Raths postulated eight needs of elementary school children. In this study an attempt was made to find out whether elementary school children felt a need for any of the eight postulated needs. Among the eight needs postulated in the earlier study was the need for achievement.

It is possible that adults have these same needs, although they would probably be expressed differently. For the in-service teacher, achievement takes the form of success in her efforts to promote learning. It would, therefore, seem that an in-service training program should help a teacher to become more successful in her efforts to promote learning. For purposes of more clearly directed research, the need for achievement will be broken down into more specifically expressed needs. This study will attempt to find out to what extent teachers express desire for training in any of the areas of postulated achievement needs and to determine the relationship between training needs expressed by teachers and the training experiences the college instructors and school administrators feel the teachers need.

### I. Statement of the Problem

The purpose of the present study is two-fold: first, to develop a valid and reliable instrument for determining in-service training needs of teachers and, secondly, to test the following relational hypotheses:

1. Training needs as expressed by the teachers participating in the Atlanta Area Teacher Education Service are related to the training experiences their college instructors believe the tea-

chers need.

2. Training needs as expressed by Atlanta Area Teacher Education Service teachers are related to training experiences their supervisors and administrators believe the teachers need.

3. The various experiences expressed as training needs by the teachers participating in the Atlanta Area Teacher Education Service are interrelated. That is, a need for greater understanding of social problems may be related to the need for a clarified educational philosophy.

4. The various training needs expressed by the Atlanta Area Teacher Education Service teachers are related to such factors as: age, sex, undergraduate major, subjects taught, etc.

Several sub-hypotheses implied but not stated here will be developed and discussed fully in Section VIII.

## II. Need for the Study

By the end of its third year of operation the Atlanta Area Teacher Education Service, representing the combined resources of six institutions of higher learning, three county systems, and three city school systems, was serving as an in-service training agency for approximately 1000 teachers in the Atlanta area. As previously stated, the major objective of the program was to improve public school instruction by helping teachers with their on-the-job problems. How well was the Atlanta Area Teacher Education Service doing this job?

In the spring of 1948, the Coordinator and other administrators of the Atlanta Area Teacher Education Service felt the program had been in operation long enough to warrant an evaluation in terms of how well it was achieving the stated objectives. With the approval of institutions and school systems concerned, the present evaluation study was undertaken.

## III. Definitions of Terms Used

Many words frequently used in educational writings are subject to some variations in meaning. In order to avoid, insofar as possible, misunderstandings and misinterpretations, certain terms will be defined according to their intended meaning in this study.

"In-service training" is that on-the-job experience which results in the teacher's professional growth.

"Institutional" refers to institutions of higher learning.

"Cooperative In-Service Training Program" denotes a plan of cooperation between college representatives and county and city school administrators.

The working concept of educational or training need as herein used is both personal and social in reference. It incorporates both what the learner presently desires and what the learner should desirably become.

To determine the educational needs of teachers does not mean to tabulate their present wishes, desires, and inclinations, nor does it mean to point out deficiencies in their particular knowledges, skills, and attitudes. As herein used, the term "training needs" refers to clues which aid in guiding the teacher's professional growth. Since needs are peculiar to time, place, and circumstance and to the individual teacher, the determining of needs must be concerned with a specific situation. Therefore, the population for the present study should desirably be limited to teachers participating in the Atlanta Area Teacher Education Service.

# SECTION III

# REVIEW OF RELATED RESEARCH

Significant studies of cooperative in-service training of teachers before 1938. In an unpublished thesis, "The Improvement of Teachers in Service," L. T. Johnson (17) reported the results of a questionnaire study of public schools and institutions of higher learning to discover methods used for the improvement of teachers in service. He concluded that teachers are not thoroughly prepared for teaching after graduation, but should continue their professional growth while on the job. Johnson felt also that teacher-training institutions are coming to realize that their responsibilities for a teacher's education do not end with graduation, but they must continue to help the teacher in service. He recommends that the superintendents of small school systems cooperate with teacher-training institutions in organizing definite programs for in-service teacher training.

V. C. Martin (21), in 1936, conducted a survey of various state programs of in-service training of agricultural teachers. In-service activities reported by most of the states were: (1) student conferences, (2) district conferences, (3) summer school, (4) supervisory visits, (5) special teaching materials, and (6) teacher-trainer visits.

The following recommendations for improving institutional in-service training were made:

1. Increased personnel of training staff
2. More special teaching material
3. Cooperative planning by supervisors and teacher trainers
4. Graduate credit for courses
5. Summer school

Probably the most significant study of in-service teacher training of this period was a disser-

tation by H. O. Reid (26). One stated purpose of this investigation was to evaluate a cooperative institutional in-service training program by special techniques devised for the purpose. Another goal for Reid's study was to make practical recommendations for an in-service training program which might be applied to the University of Nebraska.

H. O. Reid's special techniques were Teacher Problem Check Sheets, Rating Scales, and Lists. The Problem Check Sheets were designed to get teachers to indicate their problems over a two-year period. The Rating Scales and Check Lists were used by both the teacher and superintendent to evaluate the in-service training program. Reid concluded that there is a trend toward more cooperation on the part of teacher-training institutions in developing effective programs of in-service training in public schools. He recommended that other colleges working on in-service training programs make similar studies of their own graduates' needs, using a wider field and data collected in sufficient amounts and kinds to prove adequate for reliable conclusions.

Later studies of cooperative in-service teacher education. Early in 1938, the Commission on Teacher Education was established by the American Council on Education with the task of stimulating institutions of varied types to join forces for the improvement of teacher education. Associated centers interested in the project were selected in June, 1939, and the school years of 1939-1940, 1940-1941, and 1941-1942 were designated for carrying out the cooperative study.

Various phases of the work of the Commission on Teacher Education have been reported from time to time in the American Council on Education Publications:

The Improvement of Teacher Education, a final report of the Commission on Teacher Education, American Council on Education, Washington, D.C., 1946.

State Programs for the Improvement of Teacher Education, C. E. Prall, American Council on Education, Washington, D.C., 1946.

Teacher Education in Service, C. E. Prall and C. L. Cushman, American Council on Education, Washington, D.C., 1944.

Teachers for Our Times, American Council on Education, Washington, D.C., 1944.

The aspect of the Commission's work pertinent to the present study is reported as part of Prall's (23) study of in-service training. He reported in detail the activities of three of the fourteen associated centers. These three programs will be reviewed briefly.

Teachers of the school system of Newton, Massachusetts, were asked for free and independent statements concerning the weaknesses of the existing in-service training program. These expressed concerns or needs of teachers were taken as starting points for a long-term program of improvement.

The one outstanding feature of the Houston, Texas, Center is its decentralization of the in-service training program which resulted in bringing into active participation as many as 170 different study groups. In spite of such decentralization, the central organization of the Houston study planned and carried out an ambitious three-year in-service project.

Because of its pertinence to the present study, a summary of the Houston project is quoted from Teacher Education In Service (23):

1. An inventory of present practices, special attention to those that had largely lost their effectiveness, and to additional practices which might be instituted to facilitate growth and professional improvement: teachers to be organized in temporary groups within each building for this purpose.

2. Formulation by each teacher group of a number of specific problems from the results of the inventory: teachers were expected to keep the list to matters of general concern, though school and level interests were not to be excluded.

3. Consolidation and sifting of these problems by the divisional subcommittees into compact lists for each level of the system.

4. Resubmission of the appropriate list to each school faculty for review and for the formation of permanent study groups on a voluntary basis; where there were too few teachers with a certain choice in a building to make a study group, the divisional subcommittee was to assist in forming a group with representatives from other buildings.

5. Each study group was to work as intensively as possible on its selected problem, but was to determine its own method of attack; it would receive such help as it requested. A general outline of possible methods of attack and procedure for group study was to be prepared and distributed by the general committee.

6. Upon completion of several months of intensive study, the findings on a few problems would be pooled or shared with a view to formulating reports and recommending means for their solution.

In the Des Moines, Iowa, Public Schools In-Service Training Program, considerable attention was paid to planning and working in a democratic fashion. An attempt was made to identify important problems which the teachers felt were in need of study. A sub-committee used the needs or concerns expressed by teachers to fashion an instrument to submit to all teachers for rating as to timeliness and importance of the problem. Broad areas which were thus included on the problem rating sheet were: home-school relations, community resources, understanding children, teachers' welfare, understanding objectives and procedures, and cultural and social development of the teachers.

A summary of the foregoing reviews of previously completed studies seems to show that recent emphasis in teacher training is on a long-range, cooperative, in-service training program designed, in terms of teachers' expressed training needs, to lead to personal, social and professional growth.

## SECTION IV

## METHODS OF PROCEDURE

The first step was to survey the literature of the field in an effort to locate an instrument for determining expressed educational needs of in-service teachers. Insofar as the present writer could ascertain there was no such instrument available commercially. One experimental method (25) was examined and tested in a pilot study, but was found to be impractical for purposes of the present problem. Other experimental devices and techniques were not applicable to the Atlanta Area Teacher Education Service evaluation study.

With the help of the Coordinator of the Atlanta Area Teacher Education Service and several consultants to the program, a Preferential Training Needs Record was developed. In order to locate any ambiguities and to bring to light problems of administration, early forms of the test were given to two classes of seniors and graduate students attending the University of Georgia in the Winter Quarter, 1949. The results of this test run were analyzed and the final form of the test was prepared.

In May, 1949, the Preferential Training Needs Record was administered to teachers, supervisors, administrators, and college instructors in the Atlanta Area Teacher Education Service. Section VIII will explain the results of the administration of the Preferential Training Needs Record to this group of on-the-job teachers, their administrators, and their college instructors.

### I. Sources of Data

All the basic data that have been used in the present study were gathered from the following sources:

1. Related professional literature (See Bibliography)
2. Results of Preferential Training Needs Record (see Appendix A)* administered to teachers of the Atlanta Area Teacher Education Service.
3. Results of Preferential Training Needs Record (see Appendix B) administered to college instructors, supervisors, and administrators of teachers of the Atlanta Area Teacher Education Service.
4. Personal Data Sheet (see Appendix C) of teachers, administrators, supervisors, and college instructors of the Atlanta Area Teacher Education Service.

### II. Treatment of the Data

The analyses of the present form of the test as well as items of educational personal information were made possible by transferring the data to Hollerith cards and processing the punched cards by means of IBM card sorting and tabulating machines to yield data necessary for the computation of various statistics.

Coefficients of reliability. The reliability coefficients of each of the six scales of the Preferential Training Needs Record were determined by use of the formula for Case IV as developed by Kuder and Richardson (20). This method of rational equivalence stresses the intercorrelations of the items in the test and the correlations of the items with the test as a whole. According to Garrett (10), this method of estimating the reliability of a test is superior to the split-half technique in certain aspects, but differences in reliability as found by the two methods are never very large (of the order .02).

Coefficients of validity. Kelley (18) has demonstrated that, provided one is willing to sacrifice the continuous nature of the criterion for computational convenience, the most accurate determination of item validities or internal consistencies can be obtained by comparing the upper twenty-seven and the lower twenty-seven percent of the total group. Elimination of the middle forty-six percent leads to results that are more consistent from sample to sample than those obtained from using all available cases.

* All references to Appendices may be found in original thesis on file in Library, University of Georgia.

Validity coefficients for each item were obtained by the methods described by Kelley. A table prepared by Flanagan (9) was used for estimating the values of the product-moment coefficients of correlation corresponding to given proportions of success in the upper and lower twenty-seven percent of the criterion groups.

Other statistics. All other statistics reported in this study were obtained by use of conventional procedures.

## SECTION V

## THE CONSTRUCTION OF A PREFERENTIAL TRAINING NEEDS RECORD FOR IN-SERVICE TEACHERS

From the foregoing discussion it is seen that today, as a program of instruction, in-service teacher education is very broad. It is concerned with: (1) the general cultural development of the in-service teacher; (2) his knowledge of the fields within which he is teaching, his knowledge of child growth and development, and other technical knowledges of psychological and educational nature; and (3) his skill in applying these knowledges and abilities to the problems of teaching.

### I. Areas of Postulated In-Service Training Needs

In the present study, each of these three broad aspects of in-service teacher-training was further broken down into smaller areas of training needs. These areas of training needs were the areas selected for investigation in the present study. Although the needs are interrelated and overlapping cannot be completely avoided, the following descriptions attempt to define and delimit the area of training covered by each postulated need.

Need for a philosophy by which to appraise educational purposes and practices. Every teacher needs to know why we have schools and what education should aim to accomplish, because the purpose for which schools were established and the objectives of the school must determine the policy and practice within the school. It is in the light of what education should accomplish that a teacher can decide on immediate and specific aims that ought to be followed in the daily life of the school. This principle has been called the philosophical or ultimate law of the school (15). Daily the teacher needs to know how to solve many problems ranging from the conventional "skin deep" type of "low I. Q.", "low-class home", or "laziness" to the more fundamental problem of enriching community living by means of carrying forward operations which help solve community problems. Is there a guiding principle for teachers in making such decisions?

Teachers need to know as much about child growth and development as possible. They should be familiar with their students—their homes, their fears, and their hopes. Teachers should also know a great deal about economic and social issues of the world today. Knowledge of professional techniques and subject matter is not enough. The ultimate purpose of the school furnishes the principle for deciding what shall be done in any particular situation. It is evident that the teacher's philosophy of education is extremely important. What he holds important in life and in school will affect the decisions he makes every day, and it was for this reason that the need for an educational philosophy by which to appraise educational purposes and practices was postulated as an area of training needs to be studied.

This area of in-service training is represented by the twenty items in Scale 3 of the Preferential Training Needs Record.

Need for a social philosophy by which to understand and interpret current social problems and issues. The teacher is responsible to the society he represents. This responsibility is not greatly different from the responsibility of other citizens in a democratic society, but, for the teacher, it is more serious because of his influential position. It is professionally important that the teacher have a more than ordinary knowledge and understanding of his society. He should have keen insight and knowledge into current social trends, economic problems, and political issues in order to understand adequately problems of his pupils.

He also needs competence in social understanding in order to be able to use sound judgment in making decisions which affect the school and community.

The second postulated in-service training need, therefore, is the need for a social philosophy by which to appraise current problems and issues, and is represented by the twenty items in Scale 4 of the Preferential Training Needs Record.

The need for competence in the providing of security for the learner. Frequently, learning involves so much challenge to the personality of students that, to facilitate learning, teachers need to carry on operations which contribute to feelings of security of the learner. Students must feel free to learn and to change. The teacher needs to know how to provide a permissive atmosphere where students may raise questions which bother them, how to help students acquire a knowledge of what causes our more persistent problems, and how, at times, to reassure students that answers are available even though they are not in the immediate possession of the teacher.

In other cases where open and perhaps serious conflict is impending, the teacher provides se-

curity through techniques of diversion, "saving face", and recognizing points in opposing positions.

From these situations, it is clear that the teacher's ability to provide security for the learner determines in part what and how children learn. For this reason, the third postulated training need is the need for competence in the providing of security for the learner. This area is represented by the twenty items in Scale 6 of the Preferential Training Needs Record.

The need for competence in the diagnosis and treatment of non-learning. There is almost unanimous agreement among educators and psychologists that all observable behavior is the result of causes. From this it follows that there are causes of slow learning or non-learning. The teacher of today needs to know how to identify these causes and how to relate himself to them in ways to improve learning. Teachers need to consider carefully how far their own practices are responsible for creating such conditions. They need to understand that differences in the maturity level, abilities, backgrounds, and motivation of children imply differences in the way they learn. Frequently, children do not learn because the teacher's goals are not their goals. Learners have to be purposefully involved. They have to be helped to see relationships, to plan effectively, and they must be allowed freedom to evaluate their progress. Sometimes there are environmental blocks to learning. Conditions within the home and study habits contribute to slow learning. The schedules of home and school, inadequate resources and books, rules and regulations may conflict with learning. These and countless others are the causes of non-learning.

Causes do not act singly. Because of their inter-relatedness, their relative weights are unknown. Teachers need to use these ideas as hypotheses to explore learning problems. Through their research will come greater knowledge of the cause and effect of slow or non-learning.

The fourth postulated in-service training need is the need for competence in the diagnosis and treatment of non-learning and is represented by the twenty items in Scale 2 of the Preferential Training Needs Record.

Need for competence in the eliciting of free and permissive discussions. In order to elicit free discussion and expression the classroom atmosphere must be permissive. Students must be free to express their attitudes, their values, their prejudices, and their worries. The teacher needs to recognize the right of a student to have those attitudes which seem consistent with his previous experience. Where the observer of the teacher notices restraint, repressiveness, intolerance, uncritical acceptance or uncritical rejection, these are signs of the absence of permissiveness. Where the teacher definitely elicits statements of attitudes and values, permissiveness is present.

The Commission on Teacher Education (31) pointed up the importance of free personal expression in this manner:

> For two important and by no means unrelated reasons the program of the school should encourage personal expression, nonverbal as well as verbal, in as many ways as possible. In the first place. . . it is critically important for the development of positive freedom that personal spontaneity should be encouraged. True freedom demands the expression of truth as the individual sees it, subject to rational criticism. It requires the mastery of means of expression and the acceptance of the limitations that all such means possess. The school should provide a setting in which children will find it easy to express what they think believe, but in which they will expect to have what they say or do examined in a critical though friendly fashion. Personal expression should be responsible as well as free.
>
> The second reason for the encouragement of expression is in order to improve capacity for communication and interpretation. In social existence the ability to carry thought and feeling over to others, and to comprehend the messages offered by these, is of great importance. Inability to understand or be understood is almost as great a handicap in human relations as is deliberate misinterpretation or deception. To express only what one believes to be true, and to express it clearly, must be among the aims of the free man and good citizen.

The fifth area of postulated training needs for in-service teachers is the need for competence in the creation of free and permissive discussion and is represented by the twenty items in Scale 1 of the Preferential Training Needs Record.

Need for competence in the teaching of the fundamental skills. Although today schools no longer stress the mechanical perfection of skills and memorization of facts as they formerly did, this does not mean the learner is free from the necessity of acquiring skills such as reading, writing, and typing, and of gaining information such as vocabulary, names, and places. Intelligent problem solving is not only a matter of skill and knowledge, but it also requires appropriate manipulation and the assembling of information. The point of view held here is not that skills and information should be neglected, but that they should be considered in the larger con-

texts of their use and in relation to individual personal and social development.

Since the ability to learn increases with age, the kinds of skills to be acquired and the kinds of knowledge to be mastered must be arranged according to developmental progress. That is, learning experiences should not be introduced until the child has the requisite capacity to acquire the skills and has had the requisite prior learning. The teacher needs to know how to arrange the materials of instruction in a sequence that takes into account the developmental stages of the learner.

There is little relationship between the present level of mastery of a routine skill or the ability to reproduce information and the more permanent effects of such learning. When the organization of the learning material is understood by the learner, the material is better retained. The teacher needs to make every effort to present materials, the organizational pattern of which is clear to the learner.

The last postulated area of training need is the need for competence in the teaching of fundamental skills. This area is represented by the twenty items in Scale 5 of the Preferential Training Needs Record.

## II. Purpose of the Preferential Training Needs Record

The purpose of the Preferential Training Needs Record is to explore systematically the areas of postulated educational needs described above in order to assist individual teachers in discovering the areas of needed in-service training and to help teacher-training institutions plan programs of instruction to meet these expressed training needs. It is frequent that courses of in-service training are selected because of some chance influence rather than as a result of a careful analysis of needs of the teacher concerned. For those teachers who feel they need or would like to have training in several areas, the Preferential Training Needs Record offers a way of narrowing the field of investigation.

The Preferential Training Needs Record is a means of making a methodical approach to this problem. From the scores obtained, the teacher's attention may be directed toward training areas in accordance with his own expressed needs and not as a result of chance influence or temporary training opportunity.

Specific uses of the Preferential Training Needs Record are:

1. To point out areas of training with which the teacher may not be familiar but which involve situations and activities of the type which he believes he needs to know more about.
2. To assist teacher-training institutions in planning their programs of in-service training.
3. To determine if a teacher's choice of a training program is consistent with his expressed educational needs.

## III. Construction of the Preferential Training Needs Record

The construction of the Preferential Training Needs Record was undertaken with the following basic principles in mind:

1. The instrument should explore thoroughly the six areas of postulated in-service training needs.
2. The test should yield independent scores for each area of training.
3. With slight modifications the instrument should be suitable for teachers, their administrators, and their college instructors.
4. The test should be self-administering.
5. Scoring should be mechanical, or as simple as possible.

After two years' experimentation with three different instruments the present Preferential Training Needs Record was selected as most nearly meeting the conditions stated above. The first form of the test was constructed in the following manner: Thirty statements describing problematic situations in each of the six areas of training need were prepared by two independent judges.[1] These statements were in turn examined critically by three other independent judges,[2] who made suggestions concerning the final form of each statement. At this point several methods of scaling or rating the training needs represented by the statements were considered and tried out. It was finally decided that the forced choice technique developed by G. Frederic Kuder (19) would more adequately meet the conditions laid down in the paragraph above than any other method. This technique involves the use of the triad form of item in which the examinee is asked to indicate in which of three areas he most needs training and in which of the three areas he least needs training. From the subject's choice of answers it is simple to rank the three items of the triad in order to expressed training needs.

The next step in the preparation of the Preferential Training Needs Record was to select

1. Dr. L.E.Metcalf, University of Illinois, and R.T.Osborne, University of Georgia.

2. Dr. Floyd Jordan, Atlanta Area Teacher Education Service; Dr. Harry Armegida, Brooklyn College; Dr. Louis Raths, New York University.

from the list of statements of problematic situations, prepared as described above, twenty from each area of training need; 120 items were thus selected. Each item was then chosen at random from within a scale and was matched with a randomly selected item from each of two other scales. In this manner, 40 triads were formed. These 40 triads, made up of 120 items from the six areas of in-service training, constitute the present form of the Preferential Training Needs Record in which each area is matched against each of the other five areas eight times and in which each combination of areas appears twice.

A copy of the final form of the Preferential Training Needs Record is found in Appendix A.

## IV. Administering and Scoring the Preferential Training Needs Record

Administering the Preferential Training Needs Record. The format of the test and the relatively simple directions (see Appendix A) make the Preferential Training Needs Record virtually self-administering. However, the directions should be read over with the examinee (or group) in order to be sure it is clear that only items are to be marked in each triad—one representing the area in which training is "most" needed and one representing the area in which training is "least" needed.

There is no time limit on the Preferential Training Needs Record; it takes the average teacher about 45 minutes to complete the test.

Scoring the Preferential Training Needs Record. When the subject responds to each triad of training needs in the manner indicated previously, he is, in effect, ranking the needs on a three-point scale.

For example, the fifth triad on the Preferntial Training Needs Record is:

g. The discussion method can be used for the purpose of identifying problems to be studied by the group. I find it difficult to do this. How can teachers use the discussion method as a means by which to identify problems?

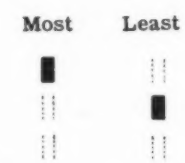
h. Children who are extremely shy or submissive may fail to learn or their learning may be lopsided. How can a teacher reduce the kind of non-learning which is caused by the excessive shyness of a child?

j. Children have many needs which are social and emotional in nature. Should the school attempt to meet these needs? Perhaps these needs should be met by some other institutions. Perhaps the school should restrict its program to the meeting of subject matter needs. Is the meeting of social and emotional needs a desirable school objective?

These items were constructed to represent the three needs: The need for competence in the eliciting of free and permissive discussions (Scale 1), the need for competence in diagnosis and treatment of non-learning (Scale 2), and the need for a social philosophy by which to understand and interpret current social problems and issues (Scale 3).

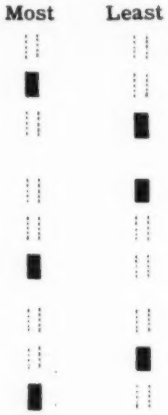
There are six possible combinations of first and last choices which may be made for each triad as follows:

1. Training represented by the first area (Scale 1) needed most and the second area (Scale 2) least. The pattern of replies is shown at the right (heavy black marks being used to indicate pencil marks). In this case, the scoring system should yield a score of 2 for Scale 1.




2. Training represented by the first area (Scale 1) needed most and the third area (Scale 3) least. Again training in the first area is preferred to both of the others and the scoring system should obtain a score of 2 for Scale 1.

3. Training represented by the second area (Scale 2) needed most and that represented by Scale 1 needed least. In this case, training in the area of Scale 1 is needed less than either of the other two, so the score should be 0 for Scale 1.




4. Training in the second area needed most and the third area needed least. In this case, training on Scale 1 is preferred to training on Scale 3 but it is not preferred to training in Scale 2. The score in this would be 1 for Scale 1.

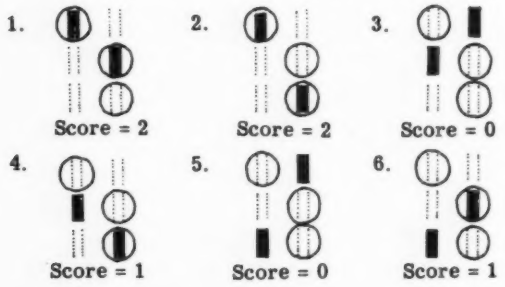
5. Training in the third area needed most and the first area needed least. This pattern shows that Scale 1 is preferred to neither of the others. The score should be 0.

6. Training in the third area needed most and in the second area least. In this case, the first area is shown to be preferred to the second but not to the third. The score should be 1.

Triads composed of any other combinations of the six scales are scored in a like manner.

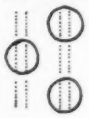
G. Frederic Kuder (19) devised a very simple scheme which obtains the desired scores when applied to each of these response patterns. It is as follows: The heavy circles around the parallel lines indicate the positions which count toward the aggregate score of the need represented by the first item of the triad.

When the composite stencil for each area of training needs is applied to the answer sheet, the six possible patterns are scored as follows:




These are the necessary scores.

When it is desired to score an area represented by the second item in the triad, the following scheme obtains the correct scores:



The area represented by the third item of the triad may be scored by the following scheme:

These three schemes are used to prepare the six composite stencils which yield aggregate scores on the six areas of training needs of the Preferential Training Needs Record.

## SECTION VI

## STANDARDIZATION OF THE PREFERENTIAL TRAINING NEEDS RECORD

### I. Reliability

In considering the reliability of the Preferential Training Needs Record, it is important to bear in mind the method used in determining this reliability. In the present study, the method of estimating the reliability by correlating the scores of two administrations of the test was necessarily ruled out because of the difficulty of scheduling the time for over 450 teachers and administrators who took part in the first administration. The method of administering an alternate form of the test was not used because the time and expense required to construct a second form of the test would have been prohibitive. The "split halves" method was not judged to be suitable for accurately estimating the reliability of this test. Therefore, the formula for Case IV as developed by Kuder and Richardson (20) was used. The

TABLE I

RELIABILITIES OF PREFERENTIAL TRAINING NEEDS RECORD SCALES
Number = 346 Teachers

| Scale | Reliability Coefficient |
|---|---|
| 1. Need for competence in eliciting permissive discussion | .85 |
| 2. Need for competence in diagnosing and treating non-learning | .86 |
| 3. Need for educational philosophy | .82 |
| 4. Need for social philosophy | .95 |
| 5. Need for competence in teaching fundamental skills | .91 |
| 6. Need for competence in providing security for learners | .89 |

TABLE II

INTERCORRELATION OF PREFERENTIAL TRAINING NEEDS RECORD SCALES
Number = 346 Teachers

| | Scale 1 | Scale 2 | Scale 3 | Scale 4 | Scale 5 | Scale 6 |
|---|---|---|---|---|---|---|
| 1. Permissive discussion | | -.29 | -.14 | -.10 | -.09 | -.28 |
| 2. Non-learning | | | -.23 | -.39 | .09 | .28 |
| 3. Educational philosophy | | | | .03 | -.24 | -.18 |
| 4. Social philosophy | | | | | -.47 | -.42 |
| 5. Fundamental skills | | | | | | -.07 |
| 6. Security giving | | | | | | — |

reliabilities obtained by this method are reported in Table I.

These r's "are slight underestimates of the true reliabilities" (19). Even so, the reliabilities for the different scales are considered satisfactorily high; the median for r's in Table I is .875.

## II. Intercorrelation

Before we discuss the information which the test gives us concerning the in-service educational needs of teachers, the general problem of the interrelationship of the areas of training need should be reviewed. In the literature it is suggested that teachers show considerable variation in their professional strengths and weaknesses and, consequently, in their in-service training needs. That is, one teacher may show one pattern of educational needs, while another of the same experience, education, and training may show an entirely different pattern. The hypothesis to be tested here is that certain in-service training needs of teachers operate relatively independently of each other. The particular needs to be studied are those postulated and identified in Section II.

The scales of the Preferential Training Needs Record developed in the present study were designed so that teachers, in responding to the items on the test, would automatically rank in order of need areas of training represented by the six scales. If the hypothesis that, in general, teachers differ in their in-service training needs is correct and if the six scales really call for these various training needs, then the scores on the separate scales should not show markedly high correlations with each other. Low correlations between scores on the different scales would be evidence that they appear to be appraising relatively specific needs or needs which do not operate together.

Table II lists the intercorrelations obtained from the base group of 346 teachers. It may be noted that the r's tend to be quite small for all scales except Scale 4, which yields r's of -.47, -.42, and -.39. The highest correlation in Table II is -.47 between need for competence in teaching fundamental skills and need for social philosophy.

## III. Validity

The writer has attempted to establish the validity of each scale by definitions; that is, to define in detail each area of training need to be tested and to construct each item in each scale in accordance with these definitions. It is for this reason that so much space in the preceding sections has been devoted to the description of the areas of postulated training needs, the statement of the purpose of the present test, and certain definitions and limitations. The selection of problematic situations for each area of educational need, the analysis of these statements by competent judges, the formulation of the statements into test questions, the process by which the questions were independently checked by a second group of judges, and the general organization of the test material have been presented in an effort to show that the test was constructed as far as possible in accordance with the definition of what the test was designed to measure. It is not the contention of the writer that this presentation satisfactorily demonstrates the validity of the test. It is hoped that, at a later date, complete statistical validation of the Preferential Training Needs Record will be possible.

The examination of the validity of items within each scale is not presented as evidence of the validity of each scale as a whole. It is, rather, evidence of the internal consistency of each scale and the degree to which each of the items on Scale 1, for example, is measuring what the whole scale purports to measure. The measure of item validity in the present study is the coefficient of correlation between the responses made to the items and total scores on the scale in which the item appears. In the next revision of the Preferential Training Needs Record items will be kept in their present form or rejected on the basis of item analysis. The item validity coefficients were determined by Kelley's method previously described. This method has been found to yield results which are extremely close approximations of the results of bi-serial correlations (10).

Item validities for the twenty items of each scale are shown in Table III.

These figures suggest that there are several items on one of the scales (Scale 2) which do not adequately discriminate between levels of training need. On four of the scales (Scales 1, 3, 5, and 6), items are fairly satisfactory, while on the remaining scale (Scale 4) only six of the items have validity coefficients below .50.

## IV. Norms

In Table IV, percentile and standard score norms are presented for the base group of 346 teachers. These norms are to be used in interpreting the Preferential Training Needs Record. The means and standard errors on the Preferential Training Needs Record scales for the various groups are reported in Section VIII.

## V. Interpretation of Scores

Scores are obtained in the following areas of in-service training need: (1) Need for greater competence in the eliciting of free and permissive discussion; (2) need for further training in

TABLE III

ITEM VALIDITY COEFFICIENTS FOR THE 420 ITEMS OF THE PREFERENTIAL TRAINING NEEDS RECORD

| Item | Scale 1 | Scale 2 | Scale 3 | Scale 4 | Scale 5 | Scale 6 |
|---|---|---|---|---|---|---|
| 1 | | .31 | | | .27 | .15 |
| 2 | .39 | | .31 | .59 | | |
| 3 | | .11 | .41 | | .28 | |
| 4 | .46 | | | .45 | | .47 |
| 5 | .40 | .30 | .39 | | | |
| 6 | | | | .62 | .41 | .18 |
| 7 | .55 | | .43 | | .21 | |
| 8 | | .39 | | .56 | | -.13 |
| 9 | .32 | | .39 | | | .34 |
| 10 | | .32 | | .55 | .42 | |
| 11 | .42 | .44 | | | | .26 |
| 12 | | | .38 | .69 | .51 | |
| 13 | | .19 | .43 | | | .43 |
| 14 | .40 | | | .64 | .48 | |
| 15 | | | .44 | .43 | | .46 |
| 16 | .43 | .40 | | | .34 | |
| 17 | | | .31 | | .36 | .41 |
| 18 | | .35 | .30 | .61 | | |
| 19 | .45 | | | | .29 | .38 |
| 20 | .17 | .32 | | .48 | | |
| 21 | | .25 | | | .39 | .17 |
| 22 | .54 | | .12 | .62 | | |
| 23 | | .40 | .54 | | .33 | |
| 24 | .24 | | | .48 | | .44 |
| 25 | .56 | .39 | .33 | | | |
| 26 | | | | .56 | .24 | .48 |
| 27 | | .07 | | .52 | | .44 |
| 28 | .53 | | .45 | | | .42 |
| 29 | | .45 | | .56 | .40 | |
| 30 | | | .25 | .45 | .33 | |
| 31 | | .13 | .35 | | | .55 |
| 32 | .41 | | | .59 | .32 | |
| 33 | | | .44 | .43 | | .16 |
| 34 | .18 | .33 | | | .37 | |
| 35 | | .46 | .33 | .66 | | |
| 36 | .28 | | | | .51 | .22 |
| 37 | .28 | .43 | | .52 | | |
| 38 | .40 | .20 | | | | .46 |
| 39 | .33 | | .34 | | .42 | |
| 40 | | | .28 | | .27 | .34 |

## TABLE IV

### PERCENTILE AND STANDARD SCORE NORMS FOR PREFERENTIAL TRAINING NEEDS RECORD
(Base Group 346 Teachers)

| R. S. | Scale 1 | | Scale 2 | | Scale 3 | | Scale 4 | | Scale 5 | | Scale 6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | SS | %ile | SS | %ile | SS | %ile | SS | %ile | SS | %ile | SS | %ile |
| 40 | | | | | | | | | | | | |
| 39 | | | | | | | 93 | 100 | 78 | 100 | | |
| 38 | | | | | | | 91 | 100 | 76 | 100 | 77 | 100 |
| 37 | | | | | | | 90 | 100 | 74 | 99 | 75 | 99 |
| 36 | | | | | | | 88 | 100 | 72 | 99 | 72 | 99 |
| 35 | 86 | 100 | | | | | 87 | 100 | 70 | 98 | 70 | 98 |
| 34 | 84 | 100 | 74 | 99 | | | 86 | 100 | 68 | 96 | 68 | 96 |
| 33 | 82 | 100 | 72 | 99 | | | 84 | 100 | 67 | 96 | 66 | 95 |
| 32 | 80 | 100 | 69 | 97 | | | 83 | 100 | 65 | 93 | 64 | 92 |
| 31 | 78 | 100 | 67 | 96 | 78 | 100 | 81 | 100 | 63 | 90 | 62 | 88 |
| 30 | 76 | 100 | 65 | 93 | 76 | 100 | 80 | 100 | 61 | 86 | 60 | 84 |
| 29 | 73 | 99 | 63 | 90 | 74 | 99 | 78 | 100 | 59 | 82 | 58 | 79 |
| 28 | 71 | 98 | 60 | 84 | 71 | 98 | 77 | 100 | 57 | 76 | 55 | 69 |
| 27 | 69 | 97 | 58 | 79 | 69 | 97 | 75 | 99 | 55 | 69 | 53 | 62 |
| 26 | 67 | 96 | 56 | 73 | 66 | 95 | 74 | 99 | 53 | 62 | 51 | 54 |
| 25 | 65 | 93 | 54 | 66 | 64 | 92 | 73 | 99 | 51 | 54 | 49 | 46 |
| 24 | 63 | 90 | 51 | 54 | 61 | 86 | 71 | 98 | 49 | 46 | 47 | 38 |
| 23 | 61 | 86 | 49 | 46 | 59 | 82 | 70 | 98 | 48 | 42 | 45 | 31 |
| 22 | 58 | 79 | 47 | 38 | 57 | 76 | 68 | 96 | 46 | 34 | 43 | 24 |
| 21 | 56 | 73 | 44 | 27 | 54 | 66 | 67 | 96 | 44 | 27 | 40 | 16 |
| 20 | 54 | 66 | 42 | 21 | 52 | 58 | 65 | 93 | 42 | 21 | 38 | 12 |
| 19 | 52 | 58 | 40 | 16 | 49 | 46 | 64 | 92 | 40 | 16 | 36 | 8 |
| 18 | 50 | 50 | 38 | 12 | 47 | 38 | 63 | 90 | 38 | 12 | 34 | 5 |
| 17 | 48 | 42 | 35 | 7 | 45 | 31 | 61 | 86 | 36 | 8 | 32 | 4 |
| 16 | 46 | 34 | 33 | 4 | 42 | 21 | 60 | 84 | 34 | 5 | 30 | 2 |
| 15 | 44 | 27 | 31 | 3 | 40 | 16 | 58 | 79 | 32 | 4 | 28 | 1 |
| 14 | 41 | 18 | 29 | 2 | 37 | 10 | 57 | 76 | 30 | 2 | 26 | 1 |
| 13 | 39 | 14 | 26 | 1 | 35 | 7 | 55 | 69 | 29 | 2 | 23 | 0 |
| 12 | 37 | 10 | 24 | 0 | 32 | 4 | 54 | 66 | 27 | 1 | 21 | 0 |
| 11 | 35 | 7 | 22 | 0 | 30 | 2 | 52 | 58 | 25 | 1 | 19 | 0 |
| 10 | 33 | 4 | 20 | 0 | 28 | 1 | 51 | 54 | 23 | 0 | 17 | 0 |
| 9 | 31 | 3 | 17 | 0 | 25 | 1 | 50 | 50 | 21 | 0 | 15 | 0 |
| 8 | 29 | 2 | | | 23 | 0 | 48 | 42 | 19 | 0 | | |
| 7 | 26 | 1 | | | | | 47 | 38 | 17 | 0 | | |
| 6 | 24 | 0 | | | | | 45 | 31 | 15 | 0 | | |
| 5 | 22 | 0 | | | | | 44 | 27 | 13 | 0 | | |
| 4 | 20 | 0 | | | | | 42 | 21 | | | | |
| 3 | | | | | | | 41 | 18 | | | | |
| 2 | | | | | | | 39 | 14 | | | | |
| 1 | | | | | | | 38 | 12 | | | | |
| 0 | | | | | | | 37 | 10 | | | | |

TABLE V

COMPARISON OF THE SCORES OF TEACHERS, ADMINISTRATORS, AND COLLEGE INSTRUCTORS ON SCALE ONE OF THE PREFERENTIAL TRAINING NEEDS RECORD

| Groups | Number | Mean | Diff. in Means | Sigma of Diff. | C. R. |
|---|---|---|---|---|---|
| Teachers - | 346 | 18.12 | | | |
| Administrators | 077 | 19.27 | -1.15 | .59 | 1.95[b] |
| Teachers - | 346 | 18.12 | | | |
| Instructors | 013 | 20.96 | -2.84 | 1.31 | 2.17[a] |
| Administrators - | 077 | 19.27 | | | |
| Instructors | 013 | 20.96 | -1.69 | 1.38 | 1.22 |

[b] Indicates that the difference is significant at the 2 percent level of confidence.
[a] Indicates that the difference is significant at the 1 percent level of confidence.

TABLE VI

COMPARISONS OF THE SCORES OF TEACHERS, ADMINISTRATORS, AND COLLEGE INSTRUCTORS ON SCALE TWO OF THE PREFERENTIAL TRAINING NEEDS RECORD

| Groups | Number | Mean | Diff. in Means | Sigma of Diff. | C. R. |
|---|---|---|---|---|---|
| Teachers - | 346 | 23.47 | | | |
| Administrators | 077 | 24.04 | - .57 | .57 | 1.00 |
| Teachers - | 346 | 23.47 | | | |
| Instructors | 013 | 22.04 | 1.43 | 1.30 | 1.10 |
| Administrators - | 077 | 24.04 | | | |
| Instructors | 013 | 22.04 | 2.00 | 1.38 | 1.45 |

No difference in this table is significant at the 5 percent level of confidence.

the diagnosis and treatment of non-learning; (3) need for further training in developing a social philosophy by which to understand and interpret current social problems and issues; (4) need for further training in developing an educational philosophy by which to appraise educational purposes and practices; (5) need for greater competence in the teaching of the fundamental skills; and (6) need for greater competence in the providing of security for the learner.

After a teacher has responded to the test and a profile has been prepared from the norm table, it is suggested that he carefully study the profile in terms of what training needs his administrators think he has and also in terms of what training his college instructors believe he needs. It is further suggested that the teacher look over the profile to see which scores are above the 75th percentile. The 75th percentile was chosen because it is a convenient point which lies between the 1% and 5% points of significance for normally distributed scores from tests having a reliability of .90 (19). Since test reliabilities vary somewhat from group to group and vary according to the method used for estimating them, there is probably little to be gained by attempting to set up separate cutting points for each scale, and there is much to be lost in convenience of interpretation. In this connection, it should be remembered that no cutting point is completely satisfactory.

If the profile contains only one high score (above 75th percentile) the educational need represented by that scale should receive careful attention from the teacher in planning his in-service training program. For example, if a teacher ranks above the 75th percentile on Scale 6, courses in the area of child growth and development, or related areas, should be considered carefully.

If the profile of a teacher contains two high scores then both areas should get consideration singly and together. It is highly improbable that a teacher will have more than two elevated scores, but if this happens, the areas represented by the high scores should be considered separately, and in combination of two. Even all the elevated areas may be considered when planning his in-service training program.

If no high scores show up on the Preferential Training Needs Record profile, scores above the 65th percentile should be inspected. Scores in this range probably have some significance but cannot be regarded with as much confidence as the higher scores. If all scores are near the median, it is possible the teacher has no outstanding educational need. Another possibility is that he may have filled in the test carelessly. Still another explanation may be that the teacher's training needs are not represented by the six postulated in-service training needs.

## SECTION VII

## EVALUATION OF THE PREFERENTIAL TRAINING NEEDS RECORD

It is the purpose of this section to describe the uses and discuss the limitations of the Preferential Training Needs Record. The writer suggests that the test which has been developed in the present study is a particularly useful instrument for determining the in-service educational needs of public school teachers for the following reasons: First, it approximates the actual classroom problems in that the teacher is asked to react to "real life" problematic situations representing six major areas of in-service training. Second, the test material is organized in such a way that the individual items are relatively stable and the scales developed from the items are relatively independent. Third, the Preferential Training Needs Record is easy to administer, score, and interpret. Fourth, the Preferential Training Needs Record may be used both as an instrument of evaluation and as a teaching device.

Approximation of the actual classroom problematic situation. Items in the several scales of the test attempt to duplicate the classroom problem by describing in detail the problematic situation and requiring the teacher to read the item thoughtfully and carefully in order to comprehend its full significance.

It was found in the pilot study for this research that the longer passages used make it possible to develop a central idea with its main points and supporting facts in a manner not possible in the shorter items frequently used in other similar tests.

The teacher is asked to respond to a wide variety of problem situations representing each of the six areas of training need postulated in the study. Thus, each area of training need corresponds to one of the six scales on the Preferential Training Needs Record, and each scale on the test is represented by twenty different problematic situations.

Informal discussions with teachers suggest that the test items catch their interest and attention and thus probably increase application and effort.

Stability of items and independence of scales of the Preferential Training Needs Record. By means of item analysis it was possible to show that 93 out of 120 items correlated .30 or better with the total scores for the various scales. Results of the study show that teachers who expressed a high degree of need for training in a particular area tended to choose all items consistently from this area as representing areas in which they needed training "most" while those teachers who scored lowest on a particular scale

tended consistently to mark all the items of this scale as items in which they needed training "least". The magnitude of the intercorrelation between the various scales would seem to indicate the scales are relatively independent. The correlation coefficients tend to be low, and twelve of fifteen are negative.

Administering, scoring, and interpreting the Preferential Training Needs Record. The Preferential Training Needs Record is self-administering and has no time limit. With only brief instructions teachers are able to read the directions and indicate their choice of items on the separate machine-scored answer sheet. Experience with over 1000 records indicates that 98 to 99 percent of the answer sheets are usable. Standard scores and percentile norms for the base group of 346 teachers have been computed, and instructions for interpreting the Preferential Training Needs Record are provided.

Uses of the Preferential Training Needs Record. The results of the Preferential Training Needs Record may be used in three ways: First, individual teachers may use the record in planning their in-service training programs. This use was described in detail in Section VI. Second, the college instructors may use it as a teaching device and as a tool for guiding further individual learning. Third, the results of the rating scale may be used by Atlanta Area Teacher Education Service administrative officials in planning course offerings in terms of the expressed educational needs of teachers.

Criticisms of the test. While the Preferential Training Needs Record seems to offer more possibilities for the determination of training needs of in-service teachers than do other tests now available, there are several points which may be criticized:

1. Certain of the items on the various scales do not discriminate well enough between teachers who express the "most" need and those who express the "least" need for training in the areas represented by the particular items.
2. Only one of the six scales appears to be sufficiently reliable to be used for individual diagnosis if we accept Kelley's standard of .94 (Scale 4, reliability coefficient of .95).
3. The six postulated areas of training need do not represent all possible areas of in-service training.
4. At the present time no system has been developed for differentiating between those teachers who respond to the Preferential Training Needs Record sincerely and those who respond with the purpose of making a good impression.

## SECTION VIII

## RESULTS OF THE EXPERIMENT

In discussing the results of the administration of the Preferential Training Needs Record to teachers participating in the Atlanta Area Teacher Education Service, public school administrators of the Atlanta Area, and the college instructors of the Atlanta Area, an effort will be made to answer the following questions:

1. Are the training needs expressed by teachers of the Atlanta Area Teacher Education Service related to the training experiences their college instructors believe the teachers need?
2. Are the training needs expressed by teachers of the Atlanta Area Teacher Education Service related to the training experiences their supervisors and administrators believe the teachers need?
3. Are the training experiences the college instructors believe the teachers of the Atlanta Area need related to the training experiences the administrators and supervisors believe the teachers need?
4. Are the various experiences expressed as training needs by the Atlanta Area Teacher Education Service teachers interrelated?
5. Are the several training needs expressed by teachers of the Atlanta Area Teacher Education Service related to variables such as age, sex, undergraduate major, subject taught, etc. ?

### I. Comparisons of the Scores of Teachers, Administrators, and College Instructors on the Six Scales of the Preferential Training Needs Record

Tables V to X, inclusive, will be discussed in an attempt to answer the first three questions raised above. It will be recalled that Scale I purports to measure need for training in the competence of eliciting free and permissive discussion. Data from Table V shows that the teachers in the present study do not agree fully with their administrators or with their college instructors as to the importance of further training in this area. Differences in mean scores on Scale 1 significant at the five percent level of confidence[3] show that both the administrators and college instructors believe, to a greater degree than do the teachers, that the teachers need further training in the eliciting of permissive discussion and expression. Although the college instructors tend to feel more strongly about this training need of teachers than do the administrators, the difference in their scores on Scale 1 is not significant.

3. Hereinafter only those differences significant at the five percent level of confidence or above will be referred to as "significant" or "statistically reliable".

TABLE VII

COMPARISONS OF THE SCORES OF TEACHERS, ADMINISTRATORS, AND COLLEGE INSTRUCTORS ON SCALE THREE OF THE PREFERENTIAL TRAINING NEEDS RECORD

| Groups | Number | Mean | Diff. in Means | Sigma of Diff. | C. R. |
|---|---|---|---|---|---|
| Teachers - | 346 | 19.21 | | | |
| Administrators | 077 | 20.22 | -1.01 | .54 | 1.87[c] |
| Teachers - | 346 | 19.21 | | | |
| Instructors | 013 | 22.35 | -3.14 | 1.05 | 2.99[a] |
| Administrators - | 077 | 20.22 | | | |
| Instructors | 013 | 22.35 | -2.13 | 1.13 | 1.88[c] |

[a] Indicates that the difference is significant at the one percent level of confidence.
[b] Indicates that the difference is significant at the five percent level of confidence.

TABLE VIII

COMPARISONS OF THE SCORES OF TEACHERS, ADMINISTRATORS, AND COLLEGE INSTRUCTORS ON SCALE FOUR OF THE PREFERENTIAL TRAINING NEEDS RECORD

| Groups | Number | Mean | Diff. in Means | Sigma of Diff. | C. R. |
|---|---|---|---|---|---|
| Teachers - | 346 | 9.48 | | | |
| Administrators | 077 | 6.59 | 2.89 | .98 | 2.94[a] |
| Teachers - | 346 | 9.48 | | | |
| Instructors | 013 | 12.15 | -2.67 | 2.88 | .93 |
| Administrators - | 077 | 6.59 | | | |
| Instructors | 013 | 12.15 | -5.56 | 2.99 | 1.86[c] |

[a] Indicates that the difference is significant at the one percent level of confidence.
[b] Indicates that the difference is significant at the five percent level of confidence.

TABLE IX

COMPARISONS OF THE SCORES OF TEACHERS, ADMINISTRATORS, AND COLLEGE INSTRUCTORS ON SCALE FIVE OF THE PREFERENTIAL TRAINING NEEDS RECORD

| Groups | Number | Mean | Diff. in Means | Sigma of Diff. | C. R. |
|---|---|---|---|---|---|
| Teachers - | 346 | 24.29 | | | |
| Administrators | 077 | 24.62 | - .33 | .70 | .47 |
| Teachers - | 346 | 24.29 | | | |
| Instructors | 013 | 18.62 | 5.67 | 2.46 | 2.30[a] |
| Administrators - | 077 | 24.62 | | | |
| Instructors | 013 | 18.62 | 6.00 | 2.52 | 2.38[a] |

[a] Indicates that the difference is significant at the one percent level of confidence.

TABLE X

COMPARISONS OF THE SCORES OF TEACHERS, ADMINISTRATORS, AND COLLEGE INSTRUCTORS ON SCALE SIX OF THE PREFERENTIAL TRAINING NEEDS RECORD

| Groups | Number | Mean | Diff. in Means | Sigma of Diff. | C. R. |
|---|---|---|---|---|---|
| Teachers - | 346 | 25.42 | | | |
| Administrators | 077 | 25.09 | .33 | .64 | .51 |
| Teachers - | 346 | 25.42 | | | |
| Instructors | 013 | 24.04 | 1.38 | 1.31 | 1.05 |
| Administrators - | 077 | 25.09 | | | |
| Instructors | 013 | 24.04 | 1.05 | 1.41 | .74 |

No difference in this table is significant at the five percent level of confidence.

TABLE XI

MEANS AND STANDARD ERRORS OF THE SIX SCALES OF THE PREFERENTIAL TRAINING NEEDS RECORD FOR TEACHERS PARTICIPATING IN THE ATLANTA AREA TEACHERS EDUCATION SERVICE: BY COUNTIES

| County | No. | | Scale 1 | Scale 2 | Scale 3 | Scale 4 | Scale 5 | Scale 6 |
|---|---|---|---|---|---|---|---|---|
| Fulton | 237 | M | 17.79 | 23.44 | 19.30 | 9.74 | 24.17 | 25.52 |
| | | SE | .33 | .29 | .29 | .51 | .36 | .30 |
| Cobb | 49 | M | 18.72 | 22.71 | 19.22 | 9.66 | 24.52 | 24.55 |
| | | SE | .65 | .71 | .59 | 1.00 | .68 | .65 |
| DeKalb | 35 | M | 18.99 | 24.09 | 19.29 | 6.64 | 24.36 | 25.57 |
| | | SE | .82 | .68 | .54 | .83 | .73 | .78 |
| Others | 28 | M | 18.64 | 24.11 | 18.21 | 10.79 | 24.93 | 24.54 |
| | | SE | .79 | .78 | .95 | 1.30 | 1.26 | 1.04 |

Intercomparisons of the mean scores by counties for each scale yielded only the following differences significant at the five percent level of confidence:

Scale 4: Fulton-DeKalb; Cobb-DeKalb; Others-DeKalb.
Scale 6: DeKalb-Cobb

Table VI, which compares the scores of teachers, administrators, and college instructors on Scale 2 of the Preferential Training Needs Record, shows these three groups tend to be in agreement as to the teachers' relative need for further training in the diagnosis and treatment of non-learning.

Scale 3 attempts to measure the need for further training in developing a philosophy of education by which to appraise educational purposes and practices. Table VII shows the extent to which teachers, administrators, and college instructors feel that teachers need training in this area. It is noted that, as on Scale 1, the three groups are not in full agreement as to the relative importance of further training in this area. Both the college instructors and administrators rank the teachers' need for greater competence in this area higher than do the teachers themselves. The college instructors believe significantly more strongly than do the administrators that the teachers need further training in developing their philosophy of education—what education should aim to accomplish and the purpose for which schools were established.

Scale 4 of the Preferential Training Needs Record purports to measure need for further training in clarifying and strengthening a social philosophy by which to understand and interpret social problems and issues. As shown in Table VIII, the teachers and college instructors appear to be in agreement as to the need of training in this area. However, the administrators differ significantly from both of the above-mentioned groups in their appraisal of the relative importance of the teachers' need for training in this area.

In Table IX, mean scores of Scale 5, which attempts to measure the teachers' need for greater competence in the teaching of fundamental skills, reveal that: (1) Teachers and administrators are in agreement as to the relative importance of further training in the area; (2) teachers and college instructors differ significantly in the relative importance they place on the teachers' further need for training in this area; (3) administrators and college instructors differ significantly in the amount of further training that they believe the teachers need in the teaching of fundamental skills; and (4) both teachers and their administrators feel more strongly than do the college instructors that the teachers need further training in the area represented by Scale 5.

Results of the administration of Scale 6 of the Preferential Training Needs Record, shown in Table X, are significant with respect to similarities among the three groups. Teachers, administrators, and college instructors tend to be in agreement as to the teachers' relative need for greater competence in the providing of security for the learner.

II. Interrelatedness of Expressed Training Needs

The intercorrelations between the various scales of the Preferential Training Needs Record were presented in Table II and were discussed in detail in Section VI, Standardization of the Preferential Training Needs Record. From this discussion, it is recalled that the scales of the Preferential Training Needs Record developed in the present study were designed so that teachers, in responding to the items on the test, would automatically rank in order of need areas of training represented by the six scales. If the hypothesis that, in general, teachers differ in their in-service training needs is correct and if the six scales really call for these various training needs, then the scores on the separate scales should not show markedly high correlations with each other. Low correlations between scores on the different scales seem to be evidence that, with the exception of Scale 4, each scale appears to be appraising relatively specific needs or needs which do not operate together.

III. Relationship Between Expressed Training Needs and Such Variables as Age, Sex, Subject Taught, etc.

Tables XI to XVII, inclusive, will be discussed in an attempt to answer the last question raised in the introduction to this section: that is, are the educational needs expressed by teachers of the Atlanta Area Teacher Education Service related to variables such as age, sex, subject taught, undergraduate major, etc.?

In these tables, 572 critical ratios were computed in order to make all necessary intercomparisons. Because of space limitations, no attempt will be made to present, interpret, or discuss all these differences in the present study. When a significant group difference was found, this fact is indicated as a footnote of the table in which the group comparisons are shown.

Table XI shows that teachers of Fulton, Cobb, and other counties who participated in this study feel the need for a better understanding of current social problems and issues, the area represented by Scale 4, more strongly than do the teachers of DeKalb County. However, DeKalb County teachers feel a greater need than do those of Cobb County for further training in the area represented by Scale 6, competence in the providing of security for the learners. Differences in mean scores for the various counties on all other scales of the Preferential Training Needs Record are not significant.

Table XII shows differences in means by sex and by age group for women teachers on the six scales of the Preferential Training Needs Record. Of the eighteen possible comparisons by age group only four yielded significant differences. The

TABLE XII

MEANS AND STANDARD ERRORS OF THE SIX SCALES OF THE PREFERENTIAL TRAINING NEEDS RECORD FOR TEACHERS PARTICIPATING IN THE ATLANTA AREA TEACHER EDUCATION SERVICE: BY SEX AND AGE GROUP

| Age and Sex | No. | | Scale 1 | Scale 2 | Scale 3 | Scale 4 | Scale 5 | Scale 6 |
|---|---|---|---|---|---|---|---|---|
| Females | | | | | | | | |
| Ages 19-33 | 55 | M | 19.25 | 22.64 | 19.09 | 9.65 | 24.85 | 24.55 |
| | | SE | .63 | .64 | .64 | .99 | .67 | .64 |
| Ages 34-48 | 172 | M | 17.24 | 23.83 | 19.13 | 9.13 | 24.62 | 25.99 |
| | | SE | .38 | .33 | .30 | .52 | .41 | .35 |
| Ages 49-63 | 34 | M | 18.91 | 23.47 | 18.53 | 9.97 | 23.15 | 25.41 |
| | | SE | .88 | .72 | .94 | 1.34 | .98 | .85 |
| All Females | 311 | M | 17.94 | 23.33 | 19.16 | 9.68 | 24.24 | 25.48 |
| | | SE | .28 | .25 | .24 | .42 | .30 | .27 |
| All Males | 35 | M | 19.10 | 24.77 | 19.71 | 7.56 | 25.04 | 25.11 |
| | | SE | .84 | .78 | .76 | .99 | .97 | .78 |

Intercomparisons of the mean scores by age group for female teachers for each scale yielded only the following differences significant at the five percent level of confidence:

Scale 1: Ages 19-33 - ages 34-48; ages 49-63 - ages 34-48
Scale 2: Ages 34-48 - ages 19-33
Scale 6: Ages 34-48 - ages 19-33

Intercomparisons of the mean scores by sex for each scale yielded only the following significant differences:

Scale 2: Males-Females
Scale 4: Females-Males

TABLE XIII

MEANS AND STANDARD ERRORS OF THE SIX SCALES OF THE PREFERENTIAL TRAINING NEEDS RECORD FOR TEACHERS PARTICIPATING IN THE ATLANTA AREA TEACHER EDUCATION SERVICE: BY GRADE TAUGHT

| Grade Taught | No. | | Scale 1 | Scale 2 | Scale 3 | Scale 4 | Scale 5 | Scale 6 |
|---|---|---|---|---|---|---|---|---|
| 1. Kindergarten-Grade 3 | 121 | M | 16.64 | 23.85 | 19.49 | 9.24 | 24.41 | 26.41 |
| | | SE | .41 | .41 | .38 | .56 | .49 | .40 |
| 2. Grades 4 and 5 | 66 | M | 18.17 | 23.91 | 19.14 | 9.26 | 23.74 | 25.18 |
| | | SE | .60 | .54 | .51 | .86 | .56 | .56 |
| 3. Grades 6 and 7 | 66 | M | 19.20 | 23.59 | 17.82 | 10.05 | 24.53 | 25.09 |
| | | SE | .56 | .45 | .48 | 1.06 | .40 | .59 |
| 4. Grades 8 thru 12 | 87 | M | 19.24 | 22.83 | 19.86 | 8.76 | 24.72 | 24.83 |
| | | SE | .56 | .50 | .51 | .73 | .60 | .52 |

Intercomparisons of the mean scores by grade taught for each scale yielded only the following differences significant at the five percent level of confidence:

Scale 1: Group 2 - Group 1; Group 3 - Group 1; Group 4 - Group 1.
Scale 3: Group 1 - Group 3; Group 2 - Group 3; Group 4 - Group 3.
Scale 6: Group 1 - Group 2; Group 1 - Group 3; Group 1 - Group 4.

TABLE XIV

MEANS AND STANDARD ERRORS OF THE SIX SCALES OF THE PREFERENTIAL TRAINING NEEDS RECORD FOR TEACHERS PARTICIPATING IN THE ATLANTA AREA TEACHER EDUCATION SERVICE: BY SUBJECTS TAUGHT

| Subject Taught | No. | | Scale 1 | Scale 2 | Scale 3 | Scale 4 | Scale 5 | Scale 6 |
|---|---|---|---|---|---|---|---|---|
| 1. Fine and Practical Arts | 18 | M | 17.72 | 23.50 | 18.83 | 8.83 | 25.50 | 26.00 |
| | | SE | 1.18 | .88 | .93 | 1.98 | 1.37 | 1.00 |
| 2. Language | 23 | M | 20.98 | 20.87 | 22.83 | 8.28 | 23.93 | 24.26 |
| | | SE | 1.10 | .74 | .75 | 1.56 | 1.30 | .98 |
| 3. Math and Science | 23 | M | 19.24 | 22.96 | 20.48 | 8.80 | 24.80 | 24.13 |
| | | SE | 1.12 | 1.09 | .96 | 1.39 | 1.06 | .97 |
| 4. Social Studies | 16 | M | 18.75 | 23.06 | 20.25 | 10.25 | 24.50 | 23.25 |
| | | SE | 1.40 | 1.36 | .71 | 2.14 | 1.14 | 1.30 |
| 5. More than one subject (includes grade teachers) | 226 | M | 17.68 | 23.77 | 19.01 | 0.69 | 24.28 | 24.17 |
| | | SE | .32 | .29 | .26 | .50 | .35 | .33 |

Intercomparisons of the mean scores by subjects taught for each scale yielded only the following differences significant at the five percent level of confidence:

Scale 1: Group 2 - Group 1; Group 2 - Group 3; Group 2 - Group 5.
Scale 2: Group 1 - Group 2; Group 5 - Group 2.
Scale 3: Group 4 - Group 5; Group 2 - Group 1; Group 2 - Group 3; Group 2 - Group 4; Group 2 - Group 5.
Scale 6: Group 1 - Group 4; Group 1 - Group 5.

TABLE XV

MEANS AND STANDARD ERRORS OF THE SIX SCALES OF THE PREFERENTIAL TRAINING NEEDS RECORD FOR TEACHERS PARTICIPATING IN THE ATLANTA AREA TEACHER EDUCATION SERVICE: BY COLLEGE FROM WHICH DEGREE WAS RECEIVED

| Undergraduate College | No. | | Scale 1 | Scale 2 | Scale 3 | Scale 4 | Scale 5 | Scale 6 |
|---|---|---|---|---|---|---|---|---|
| 1. University System | 122 | M | 17.73 | 23.68 | 18.49 | 9.40 | 24.48 | 25.97 |
| | | SE | .47 | .40 | .39 | .72 | .48 | .41 |
| 2. Private College-Georgia | 86 | M | 18.29 | 23.41 | 19.57 | 9.50 | 24.10 | 25.33 |
| | | SE | .74 | .45 | .45 | .72 | .53 | .54 |
| 3. Out-of-State College | 42 | M | 17.98 | 23.21 | 20.43 | 8.17 | 25.21 | 25.57 |
| | | SE | .75 | .78 | .72 | .98 | 1.03 | .76 |
| 4. Has not received degree | 9 | M | 16.61 | 24.33 | 20.33 | 12.17 | 23.28 | 25.67 |
| | | SE | 1.05 | 1.52 | 1.23 | 2.35 | .66 | 1.71 |

Intercomparisons of the mean scores by college from which degree was received were made for each scale and only the following differences were found to be significant at the five percent level of confidence:

Scale: Group 2 - Group 1; Group 3 - Group 1.

younger and older age groups recognize strongly the need for further training in the diagnosis and treatment of non-learning, while this training need is relatively less important to teachers in the middle age group.

Both the men and women teachers are in agreement as to the relative importance of further training in four of the six areas explored by the Preferential Training Needs Record. They differ on their ranking of Scales 2 and 4. Men teachers of the Atlanta Area Teacher Education Service feel the need for further knowledge and skill in diagnosing and treating non-learning more strongly than do women teachers. On the other hand, women express a greater need for developing a social philosophy by which to understand and interpret social problems and issues.

Table XIII yields two interesting patterns of differences. First, those teachers who teach in kindergarten, first, second, and third grades feel significantly more strongly than do teachers of any other grade level that they need greater competence in the providing of security for the learner. Second, this same group of primary teachers is less concerned about further training for increasing their competence in the eliciting of free and permissive discussions than teachers of any other grade level. These apparently conflicting findings suggest hypotheses which probably should be investigated but it is not within the scope of this research to do so.

When the teachers of the Atlanta Area Teacher Education Service are divided into groups by subject taught, as is shown in Table XIV, some of the groups are relatively small but still large enough to show interesting relationships. For example, the language teachers, more than any other group studied, recognize the need for further training in the eliciting of free and permissive discussions and in the developing of an educational philosophy. All groups expressed about the same relative need for further competence in the teaching of fundamental skills and the need for greater competence in the developing of a social philosophy. No significant group differences were found for Scales 4 and 5 in Table XIV.

In Table XV, the expressed training needs of teachers of the Atlanta Area Teacher Education Service are compared by college from which the teacher received her undergraduate training. Of the 36 possible comparisons shown in Table XV only two are significant. Graduates of out-of-state schools and graduates of private Georgia colleges feel more strongly than do University of Georgia graduates that they should have a better understanding of what education should accomplish and what should be the immediate and specific aims of the school. These teachers feel a need for the clarification of their educational philosophy.

Although there were only slight differences shown when the Atlanta Area Teacher Education Service teachers were compared on the Preferential Training Needs Record by college from which they received their undergraduate training, an interesting pattern of difference is shown in Table XVI in which the same teachers are compared by undergraduate major. In this table it is seen that the foreign language teachers more than any other group of teachers recognize the need for greater competence in the diagnosis and treatment of non-learning. Once again it must be said that it is not the purpose of this research to investigate hypotheses suggested by the findings of the study. At some later date this relationship should be explored by follow-up studies with different groups of teachers and the implications for the Atlanta Area Teacher Education Service made available.

In the last table to be discussed in this section, 270 intercomparisons of the responses of participating teachers to the Preferential Training Needs Record are made. In Table XVII the participating teachers are compared with respect to courses they have taken in the Atlanta Area Teacher Education Service. It will be noted that the total number of subjects reported in this table is greater than the number of the base group. This is accounted for by the fact that some participants had taken more than one course in the Atlanta Area Teacher Education Service. As only eleven percent of the differences reported in Table XVII are significant, no consistent pattern is apparent. It might be said, however, that both the group of teachers which had had the course in appreciation of arts felt a greater need for further training in the providing of security for the learner than did the teachers who had had other courses in the Atlanta Area Teacher Education Service.

## SECTION IX

## SUMMARY, FINDINGS AND RECOMMENDATIONS, AND SUGGESTIONS FOR FURTHER RESEARCH

### I. Summary of the Study

Purposes of the study. There were two basic purposes of this study: First, to construct a valid and reliable instrument for determining the training needs of in-service teachers, and second, to administer the instrument to teachers of the Atlanta Area Teacher Education Service in an attempt to discover the expressed training needs of this group of teachers. In planning the study it was hoped that information obtained would assist the Executive Committee of the Atlanta Area Teacher Education Service in carrying out the main purpose of the program which has been stated ". . . to combine the resources of the

TABLE XVI

MEANS AND STANDARD ERRORS OF THE SIX SCALES OF THE PREFERENTIAL TRAINING NEEDS RECORD FOR TEACHERS PARTICIPATING IN THE ATLANTA AREA TEACHER EDUCATION SERVICE: BY UNDERGRADUATE MAJOR

| Undergraduate Major | No. | | Scale 1 | Scale 2 | Scale 3 | Scale 4 | Scale 5 | Scale 6 |
|---|---|---|---|---|---|---|---|---|
| 1. Agriculture, Home Economics, or Fine Arts | 19 | M | 17.50 | 24.32 | 20.53 | 9.08 | 24.66 | 24.63 |
| | | SE | .68 | 1.17 | .87 | 2.14 | 1.17 | 1.02 |
| 2. Business Administration and Education | 118 | M | 18.14 | 23.80 | 18.64 | 9.50 | 24.65 | 25.53 |
| | | SE | .47 | .40 | .38 | .61 | .45 | .46 |
| 3. Engineering, Architecture, or Science | 33 | M | 17.50 | 23.18 | 20.18 | 8.05 | 24.29 | 26.45 |
| | | SE | .90 | .80 | .79 | 1.11 | 1.05 | .79 |
| 4. English | 46 | M | 18.28 | 23.93 | 18.85 | 9.85 | 24.37 | 24.98 |
| | | SE | .74 | .51 | .67 | 1.46 | .75 | .78 |
| 5. Foreign Language | 14 | M | 18.07 | 26.14 | 21.00 | 11.21 | 23.79 | 25.07 |
| | | SE | .92 | 1.06 | 1.40 | 1.24 | 1.30 | .98 |
| 6. Social Sciences | 36 | M | 17.50 | 23.33 | 18.83 | 9.83 | 24.72 | 25.83 |
| | | SE | .90 | .75 | .74 | 1.22 | .97 | .85 |

Intercomparisons of the mean scores by undergraduate major were made for each scale and the results yielded only the following differences significant at the five percent level of confidence:

Scale 2: Group 5 - Group 3; Group 5 - Group 4; Group 5 - Group 6; Group 5 - Group 2.
Scale 3: Group 1 - Group 2; Group 3 - Group 2.
Scale 4: Group 5 - Group 3.

TABLE XVII

MEANS AND STANDARD ERRORS OF THE SIX SCALES OF THE PREFERENTIAL TRAINING NEEDS RECORD FOR TEACHERS PARTICIPATING IN THE ATLANTA AREA TEACHER EDUCATION SERVICE: FOR CERTAIN COURSES

| A. A. T. E. S. Course | No. | | Scale 1 | Scale 2 | Scale 3 | Scale 4 | Scale 5 | Scale 6 |
|---|---|---|---|---|---|---|---|---|
| 1. Advanced Child Study | 64 | M | 17.56 | 23.86 | 18.80 | 8.27 | 24.75 | 26.39 |
| | | SE | .62 | .43 | .49 | .90 | .57 | .63 |
| 2. Appreciation of Arts | 27 | M | 18.09 | 22.11 | 18.67 | 12.17 | 23.72 | 26.38 |
| | | SE | .87 | .98 | 1.00 | 1.60 | 1.14 | .84 |
| 3. Art Activities in Elem. Schools | 69 | M | 17.96 | 23.13 | 19.39 | 10.45 | 23.76 | 25.57 |
| | | SE | .61 | .52 | .49 | .88 | .63 | .55 |
| 4. Educational Planning and Development | 28 | M | 18.36 | 22.50 | 19.29 | 10.64 | 24.93 | 22.61 |
| | | SE | 1.02 | 1.09 | .97 | 1.79 | 1.14 | .97 |
| 5. Fundamentals of Child Study | 118 | M | 18.04 | 23.67 | 18.58 | 9.47 | 24.21 | 25.86 |
| | | SE | .46 | .39 | .37 | .73 | .45 | .47 |
| 6. Guiding Reading of Children | 32 | M | 17.63 | 21.94 | 20.06 | 11.50 | 24.13 | 24.28 |
| | | SE | .98 | .99 | .69 | 1.43 | 1.02 | .72 |

TABLE XVII (Continued)

| A. A. T. E. S. Course | No. | | Scale 1 | Scale 2 | Scale 3 | Scale 4 | Scale 5 | Scale 6 |
|---|---|---|---|---|---|---|---|---|
| 7. Problems in Audio-Visual Aids | 43 | M | 17.87 | 23.09 | 19.81 | 10.34 | 23.55 | 25.12 |
| | | SE | .87 | .81 | .73 | 1.28 | 1.01 | .56 |
| 8. Problems in Guidance | 46 | M | 18.72 | 24.39 | 19.89 | 8.20 | 25.07 | 24.46 |
| | | SE | .71 | .66 | .65 | .80 | .82 | .74 |
| 9. Problems in Teaching | 40 | M | 18.80 | 23.18 | 19.50 | 9.60 | 23.30 | 24.23 |
| | | SE | .68 | .83 | .74 | 1.10 | 1.04 | .72 |
| 10. Problems in the Teaching of Reading | 69 | M | 17.33 | 22.87 | 18.04 | 9.91 | 24.28 | 25.52 |
| | | SE | .59 | .56 | .50 | .89 | .59 | .60 |

Intercomparisons of the mean scores for certain courses were made for each scale and only the following differences were found to be significant at the five percent level of confidence:

Scale 2: Group 1 - Group 6; Group 8 - Group 2; Group 8 - Group 6; Group 8 - Group 10.
Scale 3: Group 3 - Group 10; Group 6 - Group 5; Group 8 - Group 5; Group 6 - Group 10; Group 7 - Group 10; Group 8 - Group 10.
Scale 4: Group 2 - Group 1; Group 3 - Group 1; Group 6 - Group 1; Group 2 - Group 8; Group 3 - Group 8; Group 6 - Group 8.
Scale 6: Group 1 - Group 4; Group 1 - Group 6; Group 1 - Group 8; Group 1 - Group 9; Group 2 - Group 4; Group 2 - Group 6; Group 2 - Group 8; Group 2 - Group 9; Group 3 - Group 4; Group 5 - Group 4; Group 7 - Group 4; Group 10 - Group 4; Group 5 - Group 6; Group 5 - Group 9.

institutions (of higher learning) with those of the public school systems in locating and attempting to solve problems confronted by teachers and administrators of the public school systems of the area." (32) During the course of the study certain subsidiary information was also sought concerning the patterns of training needs of teachers from different geographical sections of the area; and also an attempt was made to study expressed in-service educational needs of teachers in terms of age, sex, previous education, and experience.

Participating teachers, administrators, and college instructors. The 346 teachers and 77 administrators and supervisors participating in the present study were enrolled in the Atlanta Area Teacher Education Service during the Spring Quarter, 1949. These numbers represent the total population with the exception of absentees on the day the Preferential Training Needs Record was administered. The thirteen college instructors who participated in this study were regular staff members of the Atlanta Area Teacher Education Service at the time the study was made.

The instrument. The instrument developed for use in this study was constructed for the purpose of exploring six postulated areas of in-service training needs:

1. Need for a philosophy by which to appraise educational purposes and practices (Scale 3).
2. Need for a social philosophy by which to understand and interpret current social problems and issues (Scale 4).
3. Need for competence in the providing of security for the learner (Scale 6).
4. Need for competence in the diagnosis and treatment of non-learning (Scale 2).
5. Need for competence in the eliciting of free and permissive discussions (Scale 1).
6. Need for competence in the teaching of the fundamental skills (Scale 5).

The test was so constructed as to yield independent scores for each area of training. The procedure employed was a forced choice technique by which the examinee ranked the six areas in order of his felt need for training in these areas.

## II. Findings and Recommendations

The purpose of the Atlanta Area Teacher Education Service is to combine the resources of institutions of higher learning in the Atlanta Area with those of public school systems in this area in locating and attempting to solve problems which confront teachers and their administrators. The Atlanta Area Teacher Education Service uses the general approach of giving teachers the kind of training experiences the teachers believe they need in order that public school instruction in the Atlanta Area may be improved. It is believed the findings of this study resulting in the following recommendations make a contribution in this direction.

Finding 1:

The teachers of the Atlanta Area Teacher Education Service are not in full agreement with their administrators nor with their college instructors concerning the teacher's need for further competence in the eliciting of free and permissive discussion.

Recommendation: Each individual teacher, administrative supervisor, and college instructor should examine the findings of the present study and analyze his own responses to the Preferential Training Needs Record in terms of the stated purposes of the Atlanta Area Teacher Education Service. After the individual analysis has been completed, groups of teachers, administrators, and college instructors should meet together under the direction of a competent leader for the purpose of clarifying their thinking concerning the relative importance of further training in the area reported by Scale 1.

Finding 2:

The teachers, administrators, and college instructors of the Atlanta Area Teacher Education Service do not differ significantly in their opinions as to the teachers' relative need for further training in the diagnosis and treatment of non-learning.

Recommendation: As the groups participating in the present study are in virtual agreement with respect to the importance of additional training in the area represented by Scale 2, it is suggested that the administrators and instructors of the Atlanta Area Teacher Education Service make every effort to provide training experiences in kinds and amounts requested by the teachers in order to implement this program of in-service education.

Finding 3:

Teachers, administrators, and college instructors do not agree as to the teachers' relative need for a more clearly defined educational philosophy. Both the college instructors and administrators rank the teachers' need for greater competence in this area higher than do the teachers themselves.

Recommendation: As a unified educational philosophy is necessary for the effective implementation of the program of the Atlanta Area Tea-

cher Education Service, it is recommended that the Coordinator provide opportunity for discussion of these differences and encourage teachers, administrators, and college instructors to think about what education should aim to accomplish, because the purpose for which schools were established and the objective of the schools determine the policy and practice within the school.

Finding 4:

The college instructors and teachers do not agree with the school administrators concerning the relative importance of further training in the area represented by Scale 4 of the Preferential Training Needs Record—need for a social philosophy by which to understand and interpret current social problems and issues.

Recommendation: It is recommended that findings of the study including item analyses and individual and group profiles be made available to the participants in the study in an attempt to reach an agreement concerning the relative importance of this important area of in-service training and to carry out more effectively the stated purposes of the Atlanta Area Teacher Education Service.

Finding 5:

Teachers and school administrators of the Atlanta Area Teacher Education Service do not agree with the college instructors concerning the teachers' relative need for greater competence in the teaching of fundamental skills. It is the opinion of the teachers and administrators that the teachers need more training in this area than the college instructors think the teachers need.

Recommendation: It is suggested that the college instructors, who disagree with both teachers and administrators concerning this training need, examine the results of this study including the individual and group profiles in terms of the stated purpose of the Atlanta Area Teacher Education Service: i. e., "to combine the resources of the institutions (of higher learning) with those of public school systems in attempting to solve the problems confronted by teachers and administrators of the public school systems of the area."(32)

Finding 6:

Teachers, college instructors, and school administrators participating in the Atlanta Area Teacher Education Service are all in agreement concerning the teachers' relative need for greater competence in the providing of security for the learner.

Recommendation: As the teachers, administrators, and college instructors are in agreement concerning the importance of further training in this area, it is recommended that the administrative officials of the Atlanta Area Teacher Education Service make every effort to provide training experiences in kinds and amounts as requested by the teachers participating in the program.

Finding 7:

The various experiences expressed as training needs by the teachers of the Atlanta Area Teacher Education Service were found to be only slightly interrelated.

Recommendation: As the six areas of training represented by the Preferential Training Needs Record were found to be only slightly interrelated. It is suggested that there may be other in-service training needs not closely related to the six areas explored by this instrument.

Finding 8:

The expressed in-service training needs of participating teachers of the Atlanta Area Teacher Education Service were found to be somewhat related to such variables as age, sex, training, and experience.

Recommendation: It is suggested that the administration of the Atlanta Area Teacher Education Service attempt to provide training experiences in accordance with Finding 8, which is described in detail in Section VIII of this study.

## III. Suggestions for Further Research

Suggestions for further work to be done on the test have grown out of the criticisms mentioned above and out of questions which the attempt to interpret the results of the test have raised.

1. Several items on the various scales of the test should be revised or replaced in accordance with item validity coefficients reported in Section VI.
2. Certain scales should be lengthened and items within the scales clarified in order to make for greater reliability of the scales.
3. The test should be expanded to include areas of training need not now represented by the six scales of the Preferential Training Needs Record.
4. Items which are now scored on only one scale may be positively related to the total scores of other scales. If so, such items should be added to the scoring stencils of the other scales to which the items are related.
5. In its present form the Preferential Training Needs Record consists of a personal data sheet, a test booklet, and a separate machine-scored answer sheet. It should be possible to combine the personal data sheet and the answer sheet on one International Business Machines form, which would make for more efficient and

economical administration of the Preferential Training Needs Record.

At the present time, the writer is making necessary plans to incorporate the above-mentioned suggestions for further research in the second revision of the Preferential Training Needs Record.

## BIBLIOGRAPHY


1. Almy, H. C. and Sorenson, Herbert. "A Teacher-Rating Scale of Determined Reliability and Validity," Educational Administration and Supervision, XVI (March 1930), pp. 179-186.

2. Anderson, H. H. "Measures of Classroom Behavior as a Guide to Teacher Selection," Understanding the Child, X (February 1941), pp. 7-10.

3. Antrell, Henry. "An Inventory of Teacher Understanding as a Guide Toward Their Improvement in Service," Educational Administration and Supervision, XXXI (September 1945), pp. 359-366.

4. Barr, A. S. and Emans, Lester M. "What Qualities are Prerequisites to Success in Teaching?" The Nation's Schools, VI (September 1930), pp. 60-64.

5. Barr, A. S. and others. "The Validity of Certain Instruments Employed in the Measurement of Teaching Ability," The Measurement of Teaching Efficiency (New York: MacMillan Co., 1935). Pp. 71-141.

6. Bossing, Nelson L. Teacher-Aptitude Tests and Teacher Selection, Research in Higher Education, Bulletin No. 12 (Washington, D.C.: United States Office of Education, 1932). Pp. 117-133.

7. Davis, Robert A. and others. "The Teaching Problems of 1075 Public School Teachers," Journal of Experimental Education, IX (September 1940), pp. 41-60.

8. Di Michael, Salvatore G. "Comparative Changes in Teacher Attitudes Resulting from Courses in Mental Hygiene and Educational Guidance," Journal of Educational Research, XXXVII (May 1944), pp. 656-669.

9. Flanagan, J. C. "General Considerations in the Selection of Test Items and a Short Method of Estimating the Product-Moment Coefficient from the Data at the Tails of Distribution," Journal of Educational Psychology, XXX (1939), pp. 674-680.

10. Garrett, H. E. Statistics in Education and Psychology, Third Edition (New York: Longmans, Green and Co., 1947). 487 pp.

11. Garrison, K. C. "Comparative Responses of Teachers and Student Teachers to Various Items on the Teaching Situations Test," Elementary School Journal, XLV (February 1945), pp. 334-339.

12. Hartman, George W. "Measuring Teaching Efficiency Among College Instructors," Archives of Psychology, No. 154 (Washington, D.C.: The American Psychological Association, 1933), 45 pp.

13. Hartman, George W. "Social Attitudes and Information of American Elementary School Teachers," Journal of Educational Sociology, XI (April 1938), pp. 506-511.

14. Hartman, George W. "Some Assumed Determiners of Teaching Attitudes," Journal of Educational Research, XXXI (January 1938), pp. 357-362.

15. Henderson, Stella V. Introduction to Philosophy of Education (Chicago: University of Chicago Press, 1947). 401 pp.

16. Jayne, D. C. "A Study of the Relationship Between Teaching Procedures and Educational Outcomes," Journal of Experimental Education, XIV (December 1945), pp. 101-134.

17. Johnson, L. T. The Improvement of Teachers in Service. Unpublished thesis, Greeley State Teachers College, Greeley, Colorado, 1936.

18. Kelley, T. L. "The Selection of Upper and Lower Groups for the Validation of Test Items," Journal of Educational Psychology, XXX (1939), pp. 17-24.

19. Kuder, G. Frederic. Revised Manual for Kuder Preference Record (Chicago: Science Research Associates, 1946), 30 pp.

20. Kuder, G. Frederic, and Richardson, M. W. "The Theory of Estimation of Test Reliability," Psychometrika, II (1937), pp. 151-160.

21. Martin, V. G. Unifying the Program for

Improving the Instruction of Teachers in Service. Unpublished study, Mississippi State College, 1936.

22. Prall, C. E. State Programs for the Improvement of Teacher Education (Washington, D.C.: American Council on Education, 1946). 379 pp.

23. Prall, C. E. and Cushman, Teacher Education in Service (Washington, D.C.: American Council on Education, 1944). 405 pp.

24. Raths, Louis. The Wishing Well (Columbus, Ohio: Bureau of Educational Research, Ohio State University, 1945), 4 pp.

25. Redl, Fritz. A Method of Determining Training Needs of Teachers, Division of Child Development and Teacher Personnel, Committee on Teacher Education (Washington, D.C.: American Council on Education, 1940).

26. Reid, H. O. Cooperative Institutional In-Service Training for Teachers. Mimeographed dissertation, University of Nebraska, Lincoln, Nebraska, 1938. 158 pp.

27. Seagoe, May V. "Prediction of In-Service Success in Teaching," Journal of Educational Research, XXXIX (May 1946), pp. 658-667.

28. Smalzried, N. T. and Remmers, H. H. "A Factor Analysis of the Purdue Rating for Instructors," Journal of Educational Psychology, XXXIV (September 1943), pp. 363-367.

29. Symonds, Percival M. "How Teachers Solve Personal Problems," Journal of Educational Research, XXXVIII (May 1945), pp. 641-652.

30. Symonds, Percival M. "Problems Faced by Teachers," Journal of Educational Research, XXXV (September 1941), pp. 1-15.

31. Teachers for Our Times. (Washington, D.C.: American Council on Education, 1944). 178 pp.

32. The Atlanta Area Teacher Education Service, Bulletin No. 1 (Emory, Georgia: Emory University, 1947). 16 pp.

33. The Improvement of Teacher Education, a Final Report of the Commission of Teacher Education (Washington, D.C.: American Council on Education, 1946). 283 pp.

34. Von Eschen, D. R. "The Improvability of Teachers in Service," Journal of Experimental Education, XIV (December 1945), pp. 135-156.

# TEACHER MORALE AND CURRICULUM DEVELOPMENT: A STATISTICAL ANALYSIS OF RESPONSES TO A REACTION INVENTORY*

WILLIAM E. COFFMAN**
Oklahoma Agricultural and Mechanical College
Stillwater, Oklahoma

## SECTION I

## INTRODUCTION

Teachers are important people. They are the people directly responsible for the education of the children and youth of our country. The curriculum of the school is largely what they make it. The professor of education, the school administrator, or the curriculum director may have a large part in determining the content of printed courses of study. They may be responsible for much of the talking and writing in the field of education. But what goes on in the school depends on the teacher in the classroom—on the way he accepts and implements the ideas of the experts or adds his own creative touch based on his unique experience with a particular group of pupils. The teacher, then, is a key person in any program of curriculum development.

The whole personality of the teacher is involved in the educative process. A teacher brings to his job a background of training and experience which partly determines his behavior in his contact with children and young people. But his actions may be affected by the behavior of the people around him: the school administrator, the other teachers in the school, the children, and the members of the community in which the school is located. Too, his ways of working may be influenced by the physical environment of the school, such as the types of learning aids which are available , the size of the class, the physical condition of the building. All these influences make up the field of forces which determine the behavior of the teacher in the educational situation.

When the reaction of the teacher is overt, there is an immediate change in the environment field of other people in the situation. Sometimes, however, the reaction is emotional, changing the inner feelings of the teacher but not producing a change in the situation for other people. The accumulation of these emotional reactions over a period of time may produce a complex feeling tone associated with the job which plays a large part in determining the way a teacher acts in relation to new situations. Morale, which previously referred to the feeling tone associated with persistent effort in the face of unfavorable conditions, has come in recent years to refer to this complex feeling tone arising from a person's successive emotional reactions on the job.

Research studies of job efficiency have indicated that a person's feelings about his job are related to his rate of production. In the industrial setting, criteria for determining job efficiency are available so that the relationship can be clearly established. In the educational setting, such criteria are not readily available, but there seems to be a close enough parallel between the two for the educator to infer that the efficiency of the teacher is related to how he feels about his job.[1] Knowledge about morale and its correlates should be of use to the person who is working for

---

* Portion of a project report submitted in partial fulfillment of the requirements for the degree of Doctor of Education in the Advanced School of Education, Teachers College, Columbia University, 1949.

** Special thanks are due to Prof. Ruth Cunningham who first called attention to possibilities in the data collected by the Horace Mann-Lincoln Institute and who guided the study as a member of the project committee, to Prof. Percival Symonds who brought to the study his wide experience with problems of teacher personality, and to Prof. Robert Thorndike who in his capacity as the major advisor offered many suggestions at all stages of the investigation.

1. Daniel R. Davies, "Untapped Resources of School Boards," Teachers College Record, L (April, 1949), p. 460.

improved education.

The purpose of this study is to extend the knowledge of what factors influence teacher morale and to develop an instrument useful for estimating the morale of groups of teachers in relation to these factors.

Previous investigations. An examination of reports of previous investigations[2] indicates that while a large amount of descriptive data has been collected, nobody has succeeded in developing a satisfactory differential measuring instrument for exploring the relationships among the several components of morale in the educational situation. In almost every study, a generalized job satisfaction scale has been used as a criterion. The studies based on case material and on observations indicate that the worker's feelings about his job are related to job efficiency, but they do not indicate that these feelings are adequately measured by a scale of general job satisfaction. It is possible that workers may express general satisfaction with their job and yet have feelings about specific aspects of the situation which interfere with their efficiency.

The Ways-of-Working study. The Horace Mann-Lincoln Institute of School Experimentation at Teachers College, Columbia University, has been carrying on a broad investigation of the relationship of "ways of working" to curriculum development. One aspect of this study has been concerned with teacher morale. Two assumptions underly this aspect of the investigation. In the first place, the investigators assume that teacher morale is a critical factor in any successful program of curriculum development. If the findings of studies in industry are valid, then high morale and high job efficiency go together. Information about the factors producing morale will be helpful if one wishes to improve conditions in a school. In the second place, they assume that a method of estimating morale in a school will be useful. It will help the staff to determine changes in morale as a result of efforts to improve conditions. It will also help them to select schools in which to work.[3]

To gather information about morale, members of the staff developed a Teachers Reaction Inventory. One part of this inventory was a list of 80 questions.[4] By detailed study of teachers' responses to these questions tentative hypotheses have been formed concerning the nature and effect of teacher morale.

A statement of the project. The present study is designed to extend the findings of the preliminary study. Basically, the problem is to develop a sample more widely distributed than the Institute sample (which was limited to six school systems) and to analyze the responses in such a way that the interrelationships among the items become clear. The study attempts to answer these questions: (1) Do teachers' responses indicate that groups of items are related? (2) Can a study of the relationships among items lead to increased understanding of teacher morale and of its relationship to curriculum development? (3) Can the items be grouped into scales which will provide a satisfactory group measure of morale in a school or school system?

The design of the investigation permits a return to the original Institute sample for additional data. There are seven steps in the analysis:

1. Administer the Reaction Inventory to additional teachers in order to provide a sample more representative than the one gathered by the Institute Staff. Divide the sample into two parts of 400 cases each.
2. Distribute items into trial scales on the basis of rational judgment.
3. Using the first sample, calculate biserial correlation coefficients for each item with each trial scale.
4. Regroup the items using as criteria similarity of item content and similarity of correlation profile.
5. Using the second sample, calculate intercorrelations for the scales and for three background factors: (1) educational level, (2) years of experience, and (3) size of community.
6. Discuss teacher morale as revealed by the scales, that is, develop reference norms from the analysis of the second sample.
7. Apply the scales to the original Institute sample and relate the scores to Staff judgments of the effectiveness of curriculum development programs in the school systems.

---

2. See complete project report on file, Teachers College, Columbia, University.

3. Horace Mann - Lincoln Institute of School Experimentation, "Ways of Working for Curriculum Development," Teachers College Record, XLIX (February, 1948), pp. 343-346.

4. See Appendix 1 of complete report.

The Samples. This study is based on two stratified samples of 400 cases each (Sample 1 and Sample 2). The correlational analysis is based on a sub-sample of 340 cases (Sample 2-A) for which required background information was available.

The original Institute sample included responses from 837 teachers in six widely scattered school systems and represented 100% returns from some 50 schools. It was not, however, a satisfactory sample for this study because of the preponderance of cases from one school system. The final samples include 420 of these cases selected in such a way that not more than 15 percent of either sample came from any one school. To these were added 200 cases collected at Teachers College during the 1948 summer session and representing teachers from 20 states, 80 cases collected at Iowa State Teachers College during the 1948 summer session, and 100 cases collected from twelve miscellaneous schools in seven states.

The population represented by the samples differs in detail from the general population of teachers in the United States. It does not include teachers from small rural schools. It represents only teachers who voluntarily answered the questions. There is no information with respect to the sex of the respondents. The population is widely distributed with respect to other criteria of classification—school system represented, educational level, experience, tenure in present school, number of teachers in the school—so that possible biases may be investigated. Tables describing the samples are included in the complete project report.

## SECTION II

## BUILDING THE SCALES

The Reaction Inventory presents to teachers a list of 80 questions to which they may respond "YES", "NO", or "?". An analysis of the responses of a group of teachers to the Inventory is complicated and time-consuming if each question has to be considered separately. The first hypothesis to be tested was that the items in the Reaction Inventory may be grouped into meaningful sub-scales having substantial internal consistency and low intercorrelations. If this hypothesis be sustained, a basis for efficient analysis of responses becomes available.

The problem was to discover relationships among the 80 items of the Reaction Inventory. The method chosen includes five steps: (1) Group the items into trial scales on the basis of rational judgment; (2) correlate each item with each scale; (3) regroup the items on the bases of similarity of correlation profile and similarity of item content; (4) test the new scales as measuring instruments by calculating coefficients of internal consistency and distribution statistics; and (5) test the scales for independence by calculating intercorrelations.

There were difficulties in applying the method. One difficulty was that biserial coefficients of correlation were spuriously high in those cases where the item was part of the criterion scale. These coefficients were not comparable to those obtained from criterion scales independent of the item. A formula was derived for correcting biserial coefficients for spuriousness.[5] A second difficulty was finding a method for estimating internal consistency with three response items. This necessitated an extension of Hoyt's formula.[6] A third difficulty was that the selection of the items for the final scales required subjective judgment. The validity of these judgments must be appraised in relation to the data presented in this report.

The Trial Scales. On the basis of a preliminary examination of the Inventory, a set of eleven categories were postulated and descriptions developed to guide the judges in grouping the items. The following are the categories:

I. General satisfaction with teaching as a profession
II. View of the profession of teaching
III. Adjustment to professional tasks
IV. Satisfaction with particular job placement
V. Feelings of security
VI. View of the curriculum
VII. Relations with the principal
VIII. Relations with fellow-teachers
IX. Relations with the central administration of the school system
X. Relations with pupils
XI. Relations with the community served by the school

Four judges distributed the 80 items. Three or four judges agreed in the placement of 57 items. They placed two items in Category X and two items in Category XI, but since two items do not make a reliable scale, the preliminary analysis used only nine scales containing a total of 53 items.

The mean, standard deviation, range, and coefficient of internal consistency of each of the trial scales are shown in Table I. An examina-

5. See Appendix 2 of complete report.

6. See Appendix 3 of complete report.

TABLE I

MEAN, STANDARD DEVIATION, RANGE, AND COEFFICIENT OF INTERNAL CONSISTENCY FOR TRIAL SCALES (Sample 1)

| Statistic | Trial Scales I | II | III | IV | V | VI | VII | VIII | IX |
|---|---|---|---|---|---|---|---|---|---|
| High Score | 14 | 18 | 10 | 12 | 12 | 8 | 14 | 10 | 6 |
| Low Score | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Range | 13 | 16 | 10 | 12 | 12 | 8 | 14 | 9 | 6 |
| Mean | 11.16 | 10.35 | 5.29 | 6.72 | 6.97 | 5.13 | 10.44 | 6.44 | 3.95 |
| Standard Deviation | 2.82 | 3.09 | 2.92 | 2.37 | 2.68 | 2.29 | 3.31 | 2.01 | 1.63 |
| $r_{ii}$ (Hoyt) | .64 | .48 | .62 | .34 | .53 | .54 | .77 | .34 | .38 |

TABLE II

DISTRIBUTION OF ITEMS FOLLOWING PRELIMINARY ANALYSIS

| Trial Scales | Final Scales 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Extra Items | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I | 4 | 1 | | | | | | | | 1 | 1 | 7 |
| II | 2 | 5 | | 1 | | | 1 | | | | | 9 |
| III | | | 4 | | 1 | | | | | | | 5 |
| IV | | | 2 | | 1 | | | | | | 3 | 6 |
| V | | | | 4 | | | | | | | 2 | 6 |
| VI | | | | | 4 | | | | | 1 | | 5 |
| VII | | | | | | 6 | | 1 | | | | 7 |
| VIII | | | | | | | 2 | | | 2 | 1 | 5 |
| IX | | | | | | 1 | | 1 | | | 1 | 3 |
| Extra | | 2 | 2 | 3 | | 3 | 2 | 4 | 5 | 2 | 4 | 27 |
| Total | 6 | 8 | 8 | 8 | 6 | 10 | 5 | 6 | 5 | 6 | 12 | 80 |

TABLE III

FINAL SCALE 1: GENERAL LIKING FOR TEACHING

| Item No. | Item | Key | Response* Y | ? | N |
|---|---|---|---|---|---|
| 16. | Do you feel that being a teacher gives you an opportunity to make a significant contribution to society? | Y | 91.0 | 2.7 | 6.3 |
| 30. | Do you enjoy teaching? | Y | 94.3 | 3.2 | 2.5 |
| 49. | Do you often wish you had trained for and practiced a different profession or occupation? | N | 26.5 | 3.2 | 70.3 |
| 50. | Have you within the past two years seriously considered resigning your present position to accept a position other than teaching? | N | 26.3 | 1.5 | 72.2 |
| 68. | If a promising young person came to you to ask whether it is wise to become a teacher, would you advise him to do so? | Y | 51.0 | 21.7 | 27.3 |
| 80. | If you were starting over, would you go into teaching rather than into some other field? | Y | 54.0 | 17.5 | 38.5 |

* Percentage of teachers in Sample 2 making each response

TABLE IV

FINAL SCALE 2: VIEW OF THE PROFESSION

| Item No. | Item | Key | Response* Y | ? | N |
|---|---|---|---|---|---|
| 23. | Does being a teacher cramp your style socially? | N | 14.5 | 4.0 | 81.5 |
| 61. | Do most teachers enjoy their work? | Y | 55.7 | 25.8 | 18.5 |
| 66. | Is there too much competition for recognition and advancement among teachers? | N-? | 22.0 | 18.0 | 60.0 |
| 67. | Is there a great deal of talent among teachers which is never discovered? | N-? | 73.2 | 16.3 | 10.5 |
| 70. | Does one need "pull" to get advancement in education? | N | 34.5 | 38.2 | 27.3 |
| 71. | Does teaching have less prestige than other types of work done by college graduates? | N | 65.2 | 10.3 | 25.5 |
| 73. | Do appearances and personality count more than ability in getting recognition in education? | N | 31.8 | 37.2 | 31.0 |
| 77. | Do most teachers expect to make a career of teaching? | Y-? | 37.2 | 21.8 | 41.0 |

* Percentage of teachers in Sample 2 making each response.

tion of these data reveals the following:

1. The range and standard deviation of each scale are satisfactory for measuring group differences.

2. Scales I and VII are highly skewed in the negative direction indicating little differentiation at the high end of the scale. The other scales are more symmetrical although they probably differ significantly from normality.

3. The coefficients of internal consistency range from .77 to .34 indicating varying degrees of success for the method of grouping.

4. Even in the cases where internal consistency is relatively high, there is no evidence that the items belong in that scale rather than in one of the others. There is no index of the relationship of items in one scale to those in other scales.

The trial scales, then, provide only rough criteria for the preliminary analysis; they are not satisfactory as final groupings.

The Preliminary Analysis. The purpose of this analysis was to test the validity of the grouping by judgment and to effect a re-grouping which would give more consistent measures and at the same time include more of the items. The problem involved finding answers to these questions: (1) Are the items in the trial scales statistically related? (2) Are there additional items which can be included in the scales? (3) Are there additional categories into which the items may be grouped? (4) Are the items clear or ambiguous?

Two biserial correlation coefficients were calculated for each item with each scale—one for the dichotomy, YES - NO, ?, and one for the dichotomy, NO - YES, ?. The 106 coefficients spuriously high were corrected. The 18 coefficients for each item, entered on a card with the item, provided two estimates of the relationship of the item to the trial scales.

The two correlation profiles for each item gave some indication of the degree to which the item was satisfactory for measuring a single component. In most cases the two profiles were similar, indicating that the three responses to the item were on a continuum. In these cases, the profile having the more stable coefficients was chosen for further study. In a few cases there was a wide discrepancy between the two profiles. In these cases the "?" response appeared to represent a different dimension from the YES and NO responses. These items proved unsatisfactory for use in the scales.

The process of regrouping the items was essentially one of inspection. The cards were sorted and re-sorted in an effort to discover items having similar profiles and similar content. In judging the success of the process, it is important to understand the limitations imposed by the nature of the data. Here the problem was not to select from a large pool of items those which proved highly satisfactory for building scales, but rather to group these 80 items so that analysis of responses to the Inventory might be simplified. Therefore, some items were grouped together which in a future revision, with additonal items available, might be found to belong in two or more related scales.

The low internal consistency of some of the trial scales made interpretation of the correlation profiles for the items difficult. One could not determine whether an observed correlation of an item with one of these scales indicated a relationship with the common element measured by the scale or with a specific element of one or more of the items in the scale. For example, two items might correlate .35 with Scale IX. But the internal consistency of Scale IX was so low that one had no assurance that the similarity indicated a relationship between the two items. One might have repeated the analysis after one regrouping, but with only 80 items to work with such refinement seemed excessive.

The extent of shifting of items as a result of the preliminary analysis is indicated in Table II. Trial scales are indicated by Roman numerals and final scales by Arabic numerals. Correspondence between the two sets of numerals is coincidental. Thus, Trial Scale I corresponds roughly to Final Scale 1, but Trial Scale VII is related to Final Scale 6, not to Final Scale 7.

It is apparent from Table II that as a result of the preliminary analysis, Trial Scales IV and IX no longer exist. Each of the other trial scales has lost some items and gained some. There are three new scales. The trial scales contained 53 items. The final scales contain 68 items. The items are distributed more or less evenly among the final scales, that is, no scale contains very many or very few items.

The composition of the final scales, together with the percentage of teachers making each response, are indicated in Table III - XII. (The key indicates the direction of positive scoring and the point of dichotomy on which biserial "r" was based.) It should be noted that the meaning of each scale is derived from the nature of the items included and that the names given to the scales are useful only for purposes of reference.

The Final Analysis. The purpose of this analysis was to determine whether or not the scales developed in the preliminary analysis would prove satisfactory when applied to a new sample. The analysis provides answers to these questions: (1) Do the scales possess sufficient internal consistency for purposes of group measurement? (2) Do they provide satisfactory distribution of individual scores? (3) Are they relatively independent measures? (4) Are scale scores related significantly to years of experience, size of community in which the teachers work, or educational level at which he teaches?

The data of Table XIII provide the answers to the first two questions. The range and standard deviation for each scale indicate that scores are distributed widely. Most of the scales provide relatively symmetrical distributions. Scales 1 and 6, while negatively skewed, are less skewed than their counter-parts, Scales I and VII (Table I). The coefficients of internal consistency are satisfactory with the exceptions of those for Scales 2 and 10. The coefficient for Scale 2 is almost .50 and may be accepted with reservations. Scale 10 is not satisfactory for measuring group differences. It was not included in the correlational analysis.

When these data are compared with those for the trial scales given in Table I, one finds a general improvement but not a striking one. Scale 1 is definitely better than Scale I. The distribution for Scale 2 is somewhat less skewed than that for Scale II, but the internal consistency is no better. Scale 3 provides a wider distribution than Scale III, but the internal consistency is not improved. Scales 4, 5, 6, and 7 are somewhat better than Scales V, VI, VII, and VIII. Two new scales are available, both satisfactory for measuring group differences (Scales 8 and 9). One new scale (Scale 10) has proved unsatisfactory.

The statistical relationships among the scales are indicated by the correlations in Table XIV. In general, the correlations are low, indicating a high degree of independence for the scales. There are exceptions. The correlation of .72 between Scale 6 and Scale 8 is very high since the internal consistency of each scale is less than .80. If the purpose of the scales were simply to provide a consistent measure, these two scales might be made into one, but since it is also desirable to have an indication of the relationship among logically postulated components of morale, it is helpful to have two scales.

Scale 7, too, is related to Scales 6 and 8 almost to the extent of its internal consistency. It is also related significantly to Scale 2, but it is relatively independent of the other scales.

Scale 2 probably measures little not measured by the other scales. Its chief value is that it indicates how the various components combine to provide a generalized response to the profession.

The intercorrelations of the three background factors and the scales (Table XIV) indicate little relationship between components of morale and differences in the experience of teachers, in the sizes of communities in which they teach, and in the educational level at which they teach. Since these correlations are based on rough categories rather than on exact measures, they cannot be taken as exact indices of relationship; but they do serve to indicate that relationships are not high enough to explain to any great extent the scores of teachers on the several scales.

There is a consistent low negative correlation between scale scores and educational level indicating that high school teachers tend to be less satisfied than elementary school teachers. A study of the NEA Research Bureau,[7] indicating that men teachers are less satisfied than women teachers, suggests that the correlation observed in this study may simply reflect the higher proportion of men teachers in high schools.

The only significant relationship indicated for years of experience is the one with Scale 1—General Liking for Teaching. This supports the findings of previous studies that older teachers tend to be better satisfied with their jobs than do younger teachers. But the correlation while significantly different from zero is so low that one must conclude that the tendency is relatively unimportant for explaining general job satisfaction and that years of experience—and presumably age—are not significantly related to the more specific aspects of morale measured by the other scales.

None of the coefficients associated with size of community are significantly different from zero. This finding suggests that if significant relationships exist between teacher morale and the size of the community, this sample is not broad enough to reveal them. Hoppock[8] found that teachers in communities with population over 10,000 were better satisfied than teachers in smaller communities. His sample was more representative of teachers in rural schools than is the sample in this study.

Where significant relationships exist between the scale scores and these objective indices, they are so small that they do not provide a basis for explaining observed differences.

Summary. The first hypothesis of the investigation is sustained because: (1) Sixty-two of the items have been grouped into nine sub-scales which can be named; (2) the internal consistency of each scale is high enough to provide measures of group trends; and (3) the intercorrelations of the scales indicate that several relatively independent components of morale are measured.

The scales differentiate individuals as shown by the distributions of the responses, but the sizes of the coefficients of internal consistency indicate that such differentiation is not necessarily a differentiation with respect to the common element. Relationships between scale scores

7. NEA Research Bureau, "The Teacher Looks at Personnel Administration." NEA-Research Bulletin, XXIII (December, 1945), pp. 96-147.

8. Robert Hoppock, Job Satisfaction. New York: Harper and Bros., 1935, p. 32.

TABLE V

FINAL SCALE 3: ADJUSTMENT TO TEACHING ROUTINE

| Item No. | Item | Key | Response* Y | ? | N |
|---|---|---|---|---|---|
| 2. | Do you feel the need for more friends who are not teachers? | N-? | 45.5 | 2.3 | 52.2 |
| 3. | Do you find teaching confining? (No time off from morning to afternoon) | N-? | 54.0 | 4.0 | 42.0 |
| 5. | Would you do better work if you had smaller classes? | N-? | 82.0 | 5.3 | 17.7 |
| 12. | Must you spend too much out of school time in preparation for teaching? | N-? | 35.2 | 4.8 | 60.5 |
| 19. | Are you required to do too much "busy work" such as filling in forms correcting papers, recording grades? | N-? | 44.5 | 4.3 | 51.2 |
| 21. | Does teaching require more nervous energy than should be necessary? | N-? | 65.5 | 6.0 | 28.5 |
| 28. | Do you find teaching physically tiring? | N-? | 70.7 | 3.5 | 25.8 |
| 41. | Would you be happier if you were teaching other subject matter or at another grade level? | N | 14.3 | 8.5 | 77.2 |

* Percentage of teachers in Sample 2 making each response.

TABLE VI

FINAL SCALE 4: SATISFACTION WITH FINANCIAL SECURITY

| Item No. | Item | Key | Response* Y | ? | N |
|---|---|---|---|---|---|
| 22. | Do you think you could earn more money if you were doing something other than teaching? | N-? | 55.5 | 15.5 | 29.0 |
| 27. | Does your salary permit you to live in reasonable comfort? | Y | 57.2 | 8.3 | 34.5 |
| 29. | Is the salary schedule of your school system satisfactory? | Y | 38.3 | 10.0 | 51.7 |
| 32. | Is your work worth more than the salary you receive? | N-? | 70.5 | 17.0 | 12.5 |
| 62. | Does your school system provide an adequate retirement plan? | Y | 58.0 | 8.3 | 33.7 |
| 74. | Are most teachers paid fairly for the amount and type of work they do? | Y-? | 12.8 | 10.2 | 77.0 |
| 75. | Should schools offer more to teachers in the way of medical facilities? | N-? | 73.2 | 11.8 | 15.0 |
| 78. | Have you found living arrangements difficult (housing, pleasant eating places, etc.)? | N-? | 40.3 | 3.0 | 56.7 |

* Percentage of teachers making each response.

TABLE VII

FINAL SCALE 5: SATISFACTION WITH THE CURRICULUM

| Item No. | Item | Key | Response* Y | ? | N |
|---|---|---|---|---|---|
| 11. | Must you spend too much time in meetings of faculty groups or committees? | Y-? | 35.5 | 5.8 | 58.7 |
| 13. | Is there need for major revision of the curriculum for the subject or grade you teach? | N-? | 42.5 | 7.3 | 50.2 |
| 36. | Would you want the job of your principal if it were offered to you? | N-? | 10.8 | 5.7 | 83.5 |
| 37. | Should the curriculum of your school be studied with a view to revising it? | N-? | 68.7 | 15.8 | 15.5 |
| 53. | Are you willing to put time and energy into revising the curriculum of your school? | N-? | 86.0 | 8.2 | 5.8 |
| 55. | Do you feel that the curriculum of your school is too "conservative"? | N | 27.0 | 7.0 | 66.0 |

*Percentage of teachers in Sample 2 making each response.

TABLE VIII

FINAL SCALE 6: RELATIONS WITH THE PRINCIPAL

| Item No. | Item | Key | Response* Y | ? | N |
|---|---|---|---|---|---|
| 14. | Is your principal difficult to work with? | N | 10.5 | 4.0 | 85.5 |
| 26. | Are you given sufficient assistance by supervisors and principal when you want it? | Y | 71.5 | 6.0 | 22.5 |
| 35. | Do you feel that you know where you stand in the opinion of your principal? | Y | 71.5 | 10.5 | 18.0 |
| 39. | Are there people in the central office (city or county) who have a wrong opinion about you? | N | 4.0 | 38.3 | 57.7 |
| 47. | Does your principal have a wrong opinion about you? | N | 2.8 | 26.5 | 70.7 |
| 56. | Do most principals take a genuine personal interest in the teachers in their schools? | Y | 44.2 | 26.3 | 29.5 |
| 58. | Is it easy for you to express your ideas to your principal | Y | 77.7 | 4.5 | 17.8 |
| 63. | Does your principal ever blame the teachers for matters which are his own fault? | N | 20.2 | 14.3 | 65.5 |
| 65. | In general, are principals as efficient as executives in business? | Y | 34.0 | 24.0 | 42.0 |
| 72. | If you were given the job of your principal, would you handle it as he does now? | Y | 25.8 | 31.5 | 42.7 |

* Percentage of teachers in Sample 2 making each response.

and the background factors of educational level, size of community, and years of experience are negligible.

## SECTION III

## TEACHER MORALE AS REVEALED BY THE SCALES

The demonstration of statistical consistency and independence for the scales does not of itself indicate that they are of any use for exploring the relationship between teacher morale and curriculum development. The scales provide measures. They provide a series of numbers which indicate the responses of individuals to the scales. But numbers must be interpreted. Unless meaning can be attached to them, they will be of little use to persons working for curriculum change. The second hypothesis of this investigation is that the numbers can be interpreted, that is, that responses of a broad sample of teachers to the Inventory will provide a frame of reference for studying the relationship of teacher morale to curriculum development.

The test of the hypothesis requires answers to the following questions: (1) What are the natures of the various components of teacher morale revealed by the scales? (2) What relationships exist among the various components for teachers in a broad sample? (3) What are the implications of these relationships for the curriculum worker? The final test of this hypothesis is that curriculum workers can use the interpretations in field situations. The immediate test is the reasonableness of the interpretations in the light of the data.

A Description of the Final Scales. This discussion is based on the data of Tables III - XIV. Further study of follow-up interview material still to be analyzed will undoubtedly throw additional light on the meaning of the scales.

Final Scale 1: General Liking for Teaching (Table III)

The teacher with a high score on this scale says that he enjoys his work, that he feels it is important work, that he has no regrets regarding his choice of occupation, and that he would encourage promising condidates to enter the profession. The person with a low score gives opposite responses. He does not enjoy teaching, feels that his work makes no significant contribution to society, regrets his choice of occupation, and would discourage others who think of entering the profession.

The high percentage of teachers who give favorable responses to these questions suggests that as a group, teachers like the work they are doing. But there are apparently influences at work which cause almost half of them to question whether or not they would become teachers again if they could go back and make a choice.

There is a relatively high degree of internal consistency among the responses of teachers to the questions in this scale (Table XIII), indicating that teachers who give favorable responses to one question tend to give favorable responses to the others. But the relationship is not high enough to provide a direct interpretation of individual scores. Furthermore, the negatively skewed distribution indicates that differentiation among the highly satisfied teachers is slight. This may account in part for the low intercorrelations with other scales (Table XIV), but it cannot explain completely the relative independence of the scale.

One possible explanation of the independence of the scale may be that the particular influences producing general satisfaction in one individual are not the ones producing it in other individuals. Individuals differ in their needs, interests, and attitudes. One person may find satisfaction in teaching because he likes to work with children, while another finds satisfaction primarily in pleasant contacts with fellow-teachers. One, dissatisfied primarily because of basic personality maladjustment, may reflect his dissatisfaction by complaining about pay, while another, with personal inadequacies, may attribute his dissatisfactions to the behavior of the children. Too, school situations differ. In some situations the pay may be relatively high but the principal may be a poor leader; in other situations the pay may be relatively low while the principal is an excellent leader. General satisfactions and dissatisfactions expressed by teachers may be related to the differences in the situations in which they work.

But there is another possibility. Some teachers may enjoy teaching, may feel that it is important work, and may even encourage promising young people to enter the profession, and yet their morale may be low because of conditions which they do not consider related to the questions on this scale. The teacher may enjoy teaching and yet feel frustrated because the principal sets limits to the creativity of the teacher. He may enjoy teaching and yet have low morale because he has to work at another job in the evening in order to provide enough money to give his family a decent living. When some teachers feel this way and others do not, the result would be intercorrelations near zero. An expression of general liking for teaching may not be a satisfactory index to morale.

Since most investigations of morale and its correlates have used a scale of general job satis-

faction as a criterion, one may question whether or not the findings are really indications of the relative importance of the various factors in producing high morale. The approach of this study, using differential scales, seems more fruitful.

Scale 2: View of the Profession (Table IV)

The teacher with a high score on this scale feels that teaching has a satisfactory status position in the professional hierarchy, that recognition and advancement come as a result of productive effort, that teachers as a group are professional in their outlook, that teachers like their work, and that a teacher can lead a normal life with satisfactory social contacts. The teacher with a low score feels that teaching has less prestige than other types of work done by college graduates, that advancement comes from political influence rather than from professional competence, that the profession is made up of dissatisfied opportunists rather than satisfied professional people, and that community pressures and low status constitute a social handicap.

The low internal consistency of this scale is probably a reflection of the composite nature of the postulated component rather than of the inadequacy of the questions. The teacher's response to these items will reflect not only his knowledge of profession-wide conditions as reported in published studies but also his own experience with such status factors as pay, community attitudes, relations with colleagues, and the behavior of his principal. It will also reflect his own personality structure. One cannot determine which influences are primary for an individual from the score on the scale.

In view of the low standard deviation for this scale (Table XIII) one cannot expect large differences in mean scores for schools. In general, teachers recognize that the status of the profession is low. Such differences in scores as do occur will probably reflect differences in the status of the profession in those communities. Where administrators have a high view of the profession, where pay is relatively high, and where groups of teachers are professional, the scores may be higher than where conditions are less favorable. But it is possible for a group of teachers to have high status themselves and yet understand that status for the profession as a whole is low. Very high scores for a group of teachers on this scale may indicate that the group is not aware of conditions in other communities.

Scale 3: Adjustment to Teaching Routine (Table V)

The teacher with a high score on this scale is well adjusted to the routine demands of teaching—being with groups of children for a whole day at a time, keeping records and grading papers, preparing lessons. He is satisfied with his particular placement in the school. The teacher with a low score is not adjusted to routine. He finds teaching physically tiring, wishes he had smaller classes, rebels at the amount of "busy work" required, and has no energy left at the end of the day to cultivate outside friendships.

The score on this scale undoubtedly reflects personal characteristics. A person may be one who takes the routine demands of the job as a matter of course. He may like routine work, may find groups of children stimulating, may obtain so much satisfaction from his work that he feels no need for outside contacts. Or he may be a person of such dynamic energy that he can keep going after a full day's work and have a full social life in the evenings.

But scores may also reflect the nature of the situation in which the teacher works. He may be in a school where there is some free time during the day, where an enlightened administration keeps paper work to the minimum, where classes are relatively small, and where teachers are carefully placed with due regard to their individual interests and capacities. It is differences in the situation which one would expect to show when comparisons are made between mean scores for schools since individual differences may be expected to distribute themselves more or less evenly from school to school.

The distribution of scores in the reference sample (Table XIII) indicates that teachers generally feel the pressure of the routine demands of their jobs. The mean is below the middle of the range. The percentage in Table V provide more specific evidence. Over 80 percent of the teachers feel that classes are too large. Over 70 percent find teaching physically tiring. Over 65 percent feel that teaching requires more nervous energy than should be necessary. Persons who are interested in having teachers devote energy to curriculum development should give careful attention to ways in which the routine demands of the job may be reduced.

Scale 4: Satisfaction with Financial Security (Table VI)

The teacher with a high score on this scale is satisfied with the provisions for financial security offered by his job. He feels that he receives enough pay to live in reasonable comfort, that his pay is reasonable in relation to the amount of work he does, that he could not make more money in some other job, and that provisions for retirement and for meeting doctor and hospital bills are adequate. He has been able to find satisfactory living quarters. The person with a low score on this scale feels that his pay is too low,

TABLE IX

FINAL SCALE 7: RELATIONS WITH FELLOW TEACHERS

| Item No. | Item | Key | Response* Y | ? | N |
|---|---|---|---|---|---|
| 1. | Do you have sufficient opportunity to work with other teachers on significant problems? | Y-? | 56.0 | 3.8 | 40.2 |
| 9. | In your opinion, are most teachers aware of the real problems of today's society? | Y | 36.7 | 10.8 | 52.5 |
| 17. | Are the teachers in your school interested in each other's work? | Y | 60.0 | 16.5 | 23.5 |
| 45. | Have your colleagues in teaching received more recognition for their work than you have? | N | 12.0 | 22.5 | 65.5 |
| 60. | Are most of the teachers in your school interesting people? | Y | 75.7 | 7.3 | 17.0 |

* Percentage of teachers in Sample 2 making each response.

TABLE X

FINAL SCALE 8: SATISFACTION WITH DEGREE OF RECOGNITION AND PARTICIPATION

| Item No. | Item | Key | Response* Y | ? | N |
|---|---|---|---|---|---|
| 20. | Do you receive adequate recognition for your work? | Y | 58.0 | 17.8 | 24.2 |
| 40. | Are you given an adequate opportunity to help in making decisions about what happens in your school? | Y | 68.2 | 6.8 | 25.0 |
| 46. | Are you given an opportunity to help in policy making for the city or county school system? | Y | 42.0 | 12.5 | 45.5 |
| 52. | Do you receive praise when you have done a good job? | Y | 66.0 | 11.3 | 22.7 |
| 57. | Are you given sufficient opportunity to try out your own ideas? | Y | 85.2 | 4.3 | 10.5 |
| 69. | Is it true that unless a teacher agrees with the principal he hasn't a chance for recognition or advancement? | N | 18.5 | 15.5 | 66.0 |

* Percentage of teachers in Sample 2 making each response.

TABLE XI

FINAL SCALE 9: LIKING FOR CHILDREN

| Item No. | Item | Key | Response* Y | ? | N |
|---|---|---|---|---|---|
| 6. | Do your students show you the respect you feel they should? | Y | 83.2 | 5.5 | 11.3 |
| 15. | Do you find that the children and young people you teach provide the intellectual stimulation you would like? | Y | 49.5 | 12.0 | 38.5 |
| 44. | Would you prefer to work in a school where the children have higher I. Q's? | N-? | 22.5 | 10.5 | 67.0 |
| 64. | Would you prefer to work in a school where the children are of higher economic level? | N | 13.3 | 8.0 | 78.7 |
| 76. | Does working with immature children and young people tend to narrow a teacher's interests? | N | 33.5 | 8.5 | 58.0 |

* Percentage of teachers in Sample 2 making each response.

TABLE XII

FINAL SCALE 10: DESIRE FOR SELF-IMPROVEMENT

| Item No. | Item | Key | Response* Y | ? | N |
|---|---|---|---|---|---|
| 10. | Are you a better teacher today than you were two years ago? | Y | 80.7 | 13.3 | 6.0 |
| 25. | Are many of the teachers in your school more successful in their work than you are? | N | 14.8 | 34.7 | 50.5 |
| 31. | Do you plan to be a teacher five years from now? | Y | 60.0 | 20.2 | 19.8 |
| 33. | Should you be given more in-service training for your work? | Y | 47.5 | 15.0 | 37.5 |
| 43. | Do you wish you knew more about the work of the other teachers in your school? | Y | 69.2 | 6.3 | 24.5 |
| 54. | Do you feel that the curriculum of your school is too "progressive"? | N | 9.8 | 4.5 | 85.7 |

* Percentage of teachers in Sample 2 making each response.

TABLE XIII

MEAN, STANDARD DEVIATION, RANGE, AND COEFFICIENT OF INTERNAL CONSISTENCY FOR FINAL SCALES (Sample 2)

| Statistics | Final Scale* 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| High Score | 12 | 16 | 16 | 16 | 12 | 20 | 10 | 12 | 10 | 12 |
| Low Score | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| Range | 12 | 16 | 16 | 16 | 12 | 19 | 10 | 12 | 10 | 11 |
| Mean | 9.16 | 8.44 | 7.39 | 6.43 | 5.62 | 13.94 | 6.49 | 8.39 | 7.17 | 8.82 |
| Standard Deviation | 2.90 | 2.93 | 3.60 | 3.61 | 2.60 | 4.43 | 2.49 | 3.35 | 2.44 | 2.22 |
| $r_{ii}$ (Hoyt) | .72 | .45 | .58 | .63 | .51 | .79 | .52 | .77 | .53 | .36 |

* For description of final scales, see Tables III - XII.

that he could make more money in some other job, that provisions for retirement are inadequate, and that his living arrangements are unsatisfactory. He would like help in meeting emergency doctor and hospital bills.

Scores on this scale reflect more than the actual pay received. For one thing, a score may reflect the size of the pay in relation to that of other teachers in surrounding communities or to that of other workers the teacher may know. Too, if other conditions of work are satisfactory, teachers may be less likely to complain about the financial return from their jobs. Even high pay is inadequate compensation for work in a situation where human relations are poor, where recognition for effort is lacking, or where personal freedom is limited. In addition, the expectation of the individual teacher will play a part. Single teachers living at home will usually require less pay than married men with families.

All these factors must be considered when interpreting scores on this scale. But one must not lose sight of the fact that teachers are generally dissatisfied with the financial security offered by their jobs. Seventy-seven percent of the teachers feel that teachers are not paid fairly for the amount and type of work they do. Over 50 percent think they could earn more money in some other work and that the salary schedule of their school system is inadequate. Over 34 percent say their salary does not permit them to live in reasonable comfort. The average score for teachers in the reference sample is only 6.43 where the range is from 0 to 16. (See Table XIII.)

Differences in mean scores for schools on this scale should reflect primarily differences in the situations. They will have to be referred to the actual pay of teachers in each situation, to the relative position of the pay when compared with that of other workers in the community, and to the scores for other scales if interpretations of the differences are to be valid.

Scale 5: Satisfaction with the Curriculum (Table VII)

The teacher with a high score on this scale sees no reason for changing the curriculum and is unwilling to spend time working for change. Furthermore, he does not want to become a principal, at least not in the school in which he is working. The teacher with a low score is dissatisfied with the curriculum and willing to work for change. He is probably ambitious, since he would like to be a principal.

Interpretation of scores on this scale is not easy. A high score may result from two conditions. The person with a high score may be working in a conservative school where the curriculum is inadequate when judged by advanced standards. He may be satisfied because he is able to go through the same old motions year after year without spending much time in planning. He has an easy job and likes it. He wouldn't want the job of the principal because it would require additional effort. He thinks the time spent in committee work, if any, is too much. On the other hand, he may be working in a progressive school system where, under dynamic leadership, the group has worked hard to improve the curriculum and has made significant progress. He, too, has worked hard but he is now satisfied with the results and is a little tired of it all. Why should he seek to change things further?

The low internal consistency of this scale (Table XIII) probably reflects the ambiguity of the responses. Toward the middle of the distribution various combinations are possible. For example, the teacher in a progressive school who is willing to continue working for curriculum change but who feels that little change is necessary will score in the middle of the range. Likewise the person in a conservative system who sees need for change but who is willing to "let George do it" will score in the middle of the range. In future revisions of the Inventory one might add additional items in an effort to build two scales, one measuring dissatisfaction with a conservative curriculum and one measuring willingness to work for curriculum change. At present, all one can do is to keep in mind the ambiguity of the scale and to interpret scores in relation to what is known about the school system where the measurements are obtained.

Scale 6: Relations with the Principal (Table VIII)

The person with a high score on this scale has satisfactory personal and professional relations with his principal. He knows where he stands in the opinion of his principal. He receives assistance from the principal (and from other supervisors) when he needs it. He feels that the principal is doing a good job and would handle the job the same way if he were principal. He is probably well adjusted himself, since he is not suspicious of the people who supervise his work. The person with a low score on this scale feels that the principal is a poor one who is hard to get along with and who is not helpful.

Scores on this scale may reflect the behavior of the principal or they may reflect only the kind of person making the responses. It takes two people to establish good personal relations. Where scores for groups of teachers in a school are low, one may infer that the principal is deficient as a leader since it is not likely that all of the teachers in a school will be difficult to get along with. High scores for a group of teachers should indicate that the principal knows how to get along with people. High scores do not necessarily indicate good professional leadership since teach-

ers may be quite happy working with a principal who is a "good fellow" but who never challenges them to high professional efforts.

Scale 7: Relations with Fellow-Teachers (Table IX)

The teacher with a high score on this scale finds his fellow-teachers interesting people and has sufficient opportunity to work with them on significant problems. He feels that they do not receive recognition at his expense. He looks upon them as professional people interested in moving forward together. On the other hand, the teacher with a low score feels that his fellow-teachers are not interesting people, or perhaps he just doesn't know them very well. He wishes he had more opportunity to work with other teachers on significant problems, but not necessarily with his own colleagues. He has a low opinion of the professional outlook of his fellow-teachers; he may be a little condescending toward them.

The internal consistency of this scale is relatively low, but then, this scale has only five items, each concerned with a special aspect of teacher-teacher relations. The meaning of individual scores will depend on the particular pattern of responses and on the way a teacher interprets the questions. Mean scores for teachers in a school, however, probably given an indication of the degree to which good human relations exist among the staff.

Scale 8: Satisfaction with Degree of Recognition and Participation (Table X)

The teacher with a high score on this scale feels that he receives adequate recognition for his work, that he has sufficient opportunity to participate in the planning of the policy of the school or of the school system, and that he is free to try out his own ideas. He feels that his principal will recognize merit even in a teacher who expresses opinions different from his own. The teacher with a low score feels that he receives inadequate recognition for his work, that he has no part in planning the policy of the school and of the school system, and that he must conform to the ideas of the administration if he is to get ahead.

Scores on this scale may reflect individual differences in personality. Some teachers may feel it is the duty of administrators to determine policy and to offer specific guides for the conduct of the class while others may feel they should have a part in determining the policy of the school and should be free to try out their own ideas even if they involve radical departures from usual practice. Some teachers may require only indirect expressions of recognition while others are dissatisfied unless direct expressions are given frequently. But scores may also reflect differences in the situations in which teachers are working. Where the mean scores for groups of teachers in a school or school system are high, one may infer that the system is one in which teachers have more opportunities for recognition, participation, and independent action than do teachers in schools or school systems where scores are generally low.

The relatively high internal consistency of this scale indicates that feelings of adequate recognition come when teachers have an opportunity to plan together the work of the school and when they are free to try out their own ideas. Each of these has been identified in previous studies as a correlate or morale. Apparently these aspects of high morale are closely related. It seems likely that when teachers can be included in a cooperative program of curriculum development which provides opportunity for group planning, morale may be improved. This scale may prove useful for measuring changes in the fellings of teachers which follow cooperative work.

Scale 9: Liking for Children (Table XI)

The teacher with a high score on this scale finds the children or young people in his classes stimulating, gets along well with them, and has no desire to work with children who are more gifted intellectually or better off economically. The teacher with a low score receives little stimulation from working with children or young people, has poor relations with them, and would prefer to work with brighter children and with richer children.

Here again both personal preferences and situational differences are reflected in the scores. Some teachers like children, any children; others are interested only in bright children or in children from "nice" homes. Some are working in schools where all kinds of children are present; others are working in schools where selected groups are present. Another variable may be operating. If the curriculum of the school is designed to meet the needs and interests of the pupils who come to the school, teachers should have more satisfactory relations with pupils than if the curriculum is one which pupils resist. To discover which of these influences are primary in relation to mean scores for schools, the scores will have to be interpreted in relation to the whole program of each school and to the feelings of teachers with respect to the aspects of morale measured by the other scales.

Summary. It is apparent from these descriptions that an interpretation of individual scores on these scales is difficult. In each case, a score may reflect not only factors in the situation but also the unique needs, interests, and attitudes of the teacher. An individual score will not in-

TABLE XIV

INTERCORRELATIONS OF FINAL SCALES AND BACKGROUND FACTORS, (Sample 2-A)

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | | .30 | .26 | .22 | -.02 | .21 | .12 | .28 | .33 | -.19 | .14 | -.10 |
| 2. | | | .31 | .18 | .17 | .35 | .39 | .38 | .29 | -.14 | -.03 | .09 |
| 3. | | | | .14 | .06 | .21 | .16 | .21 | .28 | -.09 | -.05 | -.08 |
| 4. | | | | | .14 | .19 | .24 | .24 | .23 | -.09 | .06 | .08 |
| 5. | | | | | | .25 | .37 | .22 | .12 | -.15 | -.05 | -.10 |
| 6. | | | | | | | .49 | .72 | .20 | -.03 | .02 | -.02 |
| 7. | | | | | | | | .46 | .19 | -.22 | .04 | .01 |
| 8. | | | | | | | | | .19 | -.10 | .03 | .04 |
| 9. | | | | | | | | | | -.10 | -.02 | -.07 |
| 10. | | | | | | | | | | | .07 | .14 |
| 11. | | | | | | | | | | | | .10 |
| 12. | | | | | | | | | | | | |

1. General Liking for Teaching
2. View of the Profession
3. Adjustment to Routine
4. Satisfaction with Financial Security
5. Satisfaction with the Curriculum
6. Relations with Principal
7. Relations with Fellow-Teachers
8. Satisfaction with Degree of Recognition and Participation
9. Liking for Children
10. Educational level (Elementary - High School)
11. Years of Teaching Experience
12. Population of School Community

TABLE XV

FACTOR LOADINGS AND COMMUNALITIES FOR THE FINAL SCALES AFTER ROTATION

| Scale No. * | Factor A | Factor B | Communality |
|---|---|---|---|
| 4 | .35 | .20 | .16 |
| 1 | .58 | .03 | .34 |
| 9 | .54 | .09 | .30 |
| 3 | .49 | .10 | .25 |
| 2 | .53 | .32 | .38 |
| 8 | .30 | .73 | .62 |
| 6 | .24 | .77 | .64 |
| 7 | .25 | .63 | .46 |
| 5 | .07 | .43 | .19 |

* For descriptions of Final Scales, see Tables III - XI.

TABLE XVI

MEAN SCORES AND DEVIATIONS OF THESE SCORES FROM THE MEAN OF SAMPLE 2 FOR TEACHERS IN TWO SCHOOL SYSTEMS

| Scale No. * | School System C (N = 45) Mean | Deviation** | School System D (N = 38) Mean | Deviation** |
|---|---|---|---|---|
| 4. | 7.91 | .41 | 5.58 | -.24 |
| 1. | 10.60 | .50 | 6.95 | -.76 |
| 9. | 7.58 | .17 | 5.50 | -.99 |
| 3. | 9.00 | .54 | 5.79 | -.45 |
| 2. | 10.20 | .58 | 8.71 | .09 |
| 8. | 9.71 | .40 | 7.68 | -.21 |
| 6. | 16.42 | .54 | 11.89 | -.46 |
| 7. | 7.42 | .37 | 6.05 | -.18 |
| 5. | 6.22 | .23 | 6.76 | .44 |

* For descriptions of the scales see Tables III - XI.

** Deviations were calculated using the following formula with parameters estimated from Sample 2: $\frac{\overline{X} - \overline{M}}{\tilde{\sigma}}$

TABLE XVII

RANKING OF SIX SCHOOL SYSTEMS ACCORDING TO MEAN SCORES ON THE SCALES AND ACCORDING TO JUDGMENTS OF THE SUCCESS OF CURRICULUM DEVELOPMENT PROGRAM

| | Ranking of School System | | | | | |
|---|---|---|---|---|---|---|
| Criterion | C | B | A | E | F | D |
| Success of Program of Curriculum Improvement. (Judgment of HMLI Staff.) | 1.5 | 1.5 | 3.5 | 3.5 | 5 | 6 |
| Scale 1 | 1 | 2 | 3 | 5 | 4 | 6 |
| Scale 2 | 1 | 2 | 6 | 5 | 3 | 4 |
| Scale 3 | 1.5 | 1.5 | 4 | 5 | 3 | 6 |
| Scale 4 | 2 | 4 | 1 | 5 | 3 | 6 |
| Scale 5* | 2 | 3 | 1 | 5 | 6 | 4 |
| Scale 6 | 1 | 2 | 4 | 3 | 5 | 6 |
| Scale 7 | 3 | 1 | 4 | 2 | 5 | 6 |
| Scale 8 | 2 | 1 | 5 | 3.5 | 3.5 | 6 |
| Scale 9 | 2 | 1 | 4 | 5 | 3 | 6 |

* The rankings on this scale have been reversed, i.e., these represent relative dissatisfaction with the curriculum.

dicate unequivocally the kind of person making the responses nor the kind of situation in which he is working. Especially when scores are in the middle of the distribution, one has no way of knowing what combination of factors are operating to produce them.

If one is comparing schools rather than individuals the interpretation should be somewhat easier. One may expect individual differences in interpretation of questions and in basic personality structure to be distributed more or less evenly among schools. Differences in mean scores should reflect differences in situations since individual variation would tend to disappear. This is not entirely true, especially if the schools being compared have few teachers. The essential value of the scale measurements will be that they indicate differences which can then be assessed by the use of clinical data.

Interrelationships among the Final Scales. In addition to knowledge about the nature of each scale, one needs information about the interrelationships among the scales for a satisfactory estimate of how various factors contribute to morale. The table of intercorrelations (Table XIV) provides information of this nature. The table is derived from scales having internal consistency coefficients ranging from .45 to .79 and so any interpretation is subject to error. There are, however, certain tentative hypotheses suggested by an analysis of the table which may be helpful pending further research.

A centroid factor analysis of the table of intercorrelations (Table XIV, variables 1-9) provides some evidence of the relationships among the scales. When two factors are extracted and the axes are rotated clockwise through an angle of 45 degrees, the factor loadings and communalities are those shown in Table XV. Relatively high loadings of Factor A are found for Scales 1, 9, 3, and 2, and of Factor B for Scales 6, 7, and 8.

An examination of the items in Scales 1, 9, 3, and 2 suggests that Factor A may be called personal affinity for teaching, that is, the general personality structure of the individual which predisposes him to like or dislike the kinds of things which make up the job of the teacher. This factor seems to account in part for scores on the scales measuring general liking for teaching adjustment to routine, liking for children, and view of the profession. To a lesser degree, it is present in scores on the scales measuring feelings of financial security, relations with the principal, relations with other teachers and feelings of recognition and participation. The negligible loading on Scale 5 may be a result of the ambiguity of this scale.

The scores on the scales included in this cluster—General Liking for Teaching, Liking for Children, and Adjustment to Routine—cannot be explained by this common factor alone. Each scale has a high degree of specificity. Individual peculiarities will operate to produce differences in scores and one would need to know much about a person to account for his responses to the questions on the scales.

The high loadings of Scales 6, 7, and 8 on Factor B suggest that this factor is concerned with human relations. In relation to the internal consistency of each scale, these loadings are high. The communalities of the scales are about as high as the internal consistencies of the scales will permit.

The content of the items suggests certain tentative conclusions. Scale 7 seems to measure the general quality of the human relations of the teacher in the school situation. The scores reflect the teacher's outgoing responses and incoming responses from other people in the situation. When group measures are obtained from this scale they probably reflect the quality of the situation which makes good human relations possible. The principal seems to be an important person so far as human relations in a school situation are concerned. He can certainly so structure the situation through personal favoritism that good human relations are impossible. It may be that the high relationship between Scale 6 and Scale 8 indicates not that teachers look to the principal, alone, for recognition, but rather that they have opportunity to experience recognition only when democratic leadership provides a setting where the qualities measured by Scale 8 can be experienced together. Interpretations of Scale 8 should also take into account the loading on the factor of personal affinity for teaching. It seems reasonable to conclude that good human relations in the situation, determined partly by the opportunities given teachers to work together under a democratic principal, are necessary for a high degree of satisfaction with the degree of recognition and participation to be felt by groups of teachers but that individual teachers may have such an affinity for teaching that they are satisfied with less than average human relations.

The factor loadings for Scale 2 reflect the interaction of personal and situational influences in determining how a teacher feels about the profession. Both of the underlying factors seem to be involved.

The small communality of Scale 4 supports previous findings that feelings about financial security are relatively independent of other feelings in the morale complex.

The ambiguity of Scale 5, Satisfaction with the Curriculum, has already been discussed. The observed correlations of Scale 5 with each of the other scales can be accounted for by the correlations with Scale 7 alone. None of the partial correlation coefficients, with Scale 7 held constant, are significantly different from zero.

Thus $r_{15.7} = -.07$, $r_{25.7} = .03$, $r_{35.7} = .00$, $r_{45.7} = .05$, $r_{56.7} = .09$, $r_{58.7} = .06$, and $r_{59.7} = .06$. This seems to indicate that good human relations are related to satisfaction with the curriculum as measured by Scale 5. But there is also a specific element measured by Scale 5. This is probably a reflection of the reaction of the teacher to the actual condition of the curriculum in the school.

It seems that scores on Scale 5 will have to be interpreted in relation to the scores on the other scales, especially Scale 7. High scores on Scale 5 are not desirable if one expects to work for curriculum change. A very high score for a group of teachers would indicate that they were satisfied with the curriculum and unwilling to work for change. On the other hand, the intercorrelations suggest that a very low score would indicate that while the teachers recognized the need for change, the human relations of the situation were not promising for successful cooperative effort. In the first case, one would have to spend some time inoculating the teachers with a mild case of dissatisfaction with the curriculum before proceeding to work for change in the curriculum. In the second case, one would certainly want to work for improved human relations in the situation before going full steam ahead with a program of curriculum change. It may be that the school offering most promise for successful curriculum development, so far as this individual scale is concerned, would be one where the teachers are neither too well-satisfied nor too dissatisfied with the curriculum.

By studying the scores of teachers on this scale in relation to scores on the other scales, one might be able to make more satisfactory judgments in the case of individual schools. One might find a school where scores were low on Scale 5 but high on Scales 6, 7, and 8. In such a school one might expect efforts to change the curriculum to be highly successful. But in a school where scores on Scale 5 were very high and scores on all of the other scales were low, one would expect little curriculum change to come from the type of program the Horace Mann-Lincoln Institute sponsors.

Summary. The study of the responses of teachers in the reference sample suggest the following as important aspects of a frame of reference for interpreting scores for groups of teachers.

1. Teachers in general like their work. For Scale 1, the mean score for the reference sample indicates relatively high liking for teaching. Expressions of liking for teaching are not by themselves a satisfactory criterion of morale.

2. Teachers feel that professional status is low. The mean score of a school on Scale 2, if significantly different from that of the reference sample, probably reflects status conditions in the situation.

3. Teachers generally feel that the routine demands of their jobs are too great. For Scale 3, the mean score for the sample indicates relative dissatisfaction in this area.

4. Teachers feel that pay is inadequate but their feelings about pay may be influenced by the comparisons they make between the pay they receive and that of other people or by the working conditions they have.

5. In general, teachers' feelings about recognition are related to their feelings about opportunities to participate in planning policy and to the freedom they have to try out their own ideas. These tend to be satisfactory when relations with the principal are satisfactory. High scores for a school probably reflect democratic practices on the part of the administration.

6. Liking for children may reflect not only the disposition of teachers and the kinds of children found in the school but also the nature of the curriculum of the school.

7. Two underlying factors seem to be involved in morale as measured by the scales: (1) personal affinity for teaching and (2) human relations. Scores on Scales 1, 2, 3, and 9 reflect to some degree the first factor. Scores on Scales 6, 7, and 8 reflect to a high degree the second factor. The principal is a key person in relation to the quality of human relations in a school.

8. The scales do not measure some important aspects related to morale. There is need for information about the relations with the central administration and relations with the community for a satisfactory interpretation of scale scores.

## SECTION IV

## APPLICATIONS

The items of the Reaction Inventory can be grouped into meaningful sub-scales suitable for measuring group differences. The responses of a broad sample of teachers provide norms to which responses of teachers in individual schools or school systems may be referred. The third hypothesis of this study is that the scales may be used to estimate the status of morale in a particular school or school system, that is, that they may provide estimates of component feelings which ultimately combine to produce the complex feeling tone called morale. This hypothesis implies that the scores of teachers in one school system will differ significantly from those of teachers in other school systems—otherwise, why measure? It implies further that these differences can be related to differences in the success of curriculum development programs in the schools. If such relationships can be demonstrated, the scales will provide useful indices for persons directing programs of curriculum im-

provement. They can never be used exclusively for providing evidence that teacher morale is satisfactory for curriculum development. The director of a program will always have to supplement the data from the scales by first hand experience in the situation and by clinical judgments based on wide experience in similar situations, but he will find the scale scores provide information more rapidly than he could gather the same information by painstaking observation.

The final test of this hypothesis is a demonstration of the usefulness of the scales in a field situation. A somewhat less satisfactory test is reported in this section. The original Institute sample provides the basis for this test. From tabulated responses it was possible to calculate mean scores for each scale for the teachers in each of the six school systems represented in the sample. Too, staff judgments of the relative success of programs of curriculum improvement in the systems were available. The analysis of these data indicate that the scales do differentiate schools and that the differences are related to the amount of curriculum change which staff members observed in the schools.

A Comparison of Two School Systems. The staff of the Institute had no difficulty in designating School System C as one in which the program of curriculum improvement was highly successful and in picking School System D as one where the program was less successful. The mean scores for the scales show differences in the components of morale for teachers in the two systems.

The mean scores are shown in Table XVI. To make the scores for the nine scales comparable, they have been translated into z-scores using as reference points the means and standard deviations for the reference population reported in Table XIII. For example, the mean for the reference population for Scale 4, as estimated from Sample 2, is 6.43 and the standard deviation is 3.61. The mean score for the teachers in School System C is 7.91. Then $\frac{7.91 - 6.43}{3.61} = .41$, indicating that the mean score for teachers in School System C is .41 standard deviations above that of the teachers in the reference population. Likewise, the mean score for teachers in School System D is .24 standard deviations below the mean score for the teachers in the reference population. The differences between the mean scores for the two school systems is .65 standard deviations. This difference is so large that it would occur less than one time in a hundred if these were random samples from the reference population.[9] It seems reasonable to conclude that the observed difference is due to something other than chance. The teachers of School System A differ from those of School System D with respect to their feelings about the financial return from their jobs.

A statistical analysis will enable one to say with confidence that an observed difference cannot be attributed to chance. It will not provide information concerning the cause of differences which are statistically significant. One may draw inferences from the data, but they must be tested by gathering other kinds of data. In the present instance, for example, one might infer that the teachers in School System C receive higher pay than those in School System D. This, however, is not the case. The pay of the teachers in School System C is much lower than that of the teachers in School System D.

Evidence concerning causes of the differences may be provided by observation of the situation. How does the pay of teachers in each school system compare with that of teachers in surrounding communities? Are working conditions in a school system so poor that teachers are dissatisfied even though pay is relatively high? Are conditions of work generally so good that teachers with personality deviations find nothing to complain about except pay?

Evidence may also be provided by scores on the other scales. The data of Table XVI reveal that as a group the teachers in School System C differ from those in School System D in their feelings about things other than financial security. The differences for Scales 1, 9, 3, 8 and 6 would occur by chance less than one time in a hundred in random samples from the same population. The differences for Scales 2 and 7 would occur less than five times in a hundred. The difference for Scale 5 is so small that it might easily have occurred by chance if these were random samples from the reference population, but in view of the significant differences for the other scales it seems more reasonable to attribute the correspondence to the ambiguity of the scale rather than to a similarity in the feelings of the two groups of teachers. For the teachers in these school systems, differences in their feelings about financial security are accompanied by differences in feelings bout the other areas measured by the scales. The mean scores for the teachers in School System C are all above the mean scores for the reference population. The mean scores, for the teachers in School System D are below the mean scores of the reference population for all of the scales except 2 and 5.

This analysis demonstrated that there are significant differences in the mean scores for eight of the nine scales for teachers in two school sys-

9. See Appendix 5 of original report.

tems—one system judged successful with respect to curriculum development and one system judged unsuccessful. The analysis offers no evidence of the relative importance of the various morale factors in relation to curriculum development. Some evidence is supplied by the analysis of the previous section, but for making highly reliable interpretations one should check a number of these profiles against the actual situations in which teachers are working.

A Comparison of Six School Systems. The ranks of six school systems represented in the original Institute sample with respect to mean scores on the nine scales and with respect to staff judgments of the success of curriculum development programs are shown in Table XVII. In general, the rankings of the staff and the rankings of the scales are similar. A rough estimate of the relative effectiveness of the scales as predictors of the judgments is given by the sums of the squares of the deviations of the scale rankings from the judgment rankings. These are: Scale 1 - 4.00, Scale 2 - 17.00, Scale - 6.50, Scale 4 - 19.00, Scale 5 - 12.00, Scale 6 - 1.00, Scale 7 - 5.00, Scale 8 - 5.00, and Scale 9 - 7.00. These figures suggest some tentative hypotheses which may be tested in future study of the scales. The data are too fragmentary to provide the basis for definite conclusions.

In the first place, the scales which provide the best measures of the two underlying factors—Scales 1, 3, and 9 and Scales 6, 7, and 8—provide better predictions of staff rankings than do the other scales. The best single predictor of staff rankings is Scale 6, Relations with the Principal. This supports the inference stated previously that the principal is a key person in a cooperative program of curriculum development. Apparently a school where the teachers have a personal affinity for teaching and where good human relations exist is a good place to work for improved curriculum. This seems obvious.

In the second place, Scale 4 is the least effective predictor. This supports the findings of previous investigators indicating that income is a minor factor in relation to the general morale of teachers. One might argue that an objective index of financial return would be more accurate as a predictor than is the scale. This is not the case. The system with the highest pay is ranked 6 by the staff. This one deviation alone would give a squared deviation of 25.00 which is more than the sum of the squared deviations for Scale 4. Neither scores on Scale 4 nor an objective index of financial return is a satisfactory predictor of the judgments.

In the third place, rankings on Scale 5 are negatively related to rankings of the judges. This means that there is a tendency for relatively low scores on Scale 5 to go with high rankings of the staff. But the rankings on Scale 5 are derived from mean scores which in every case are above the mean score for the reference population. In none of these school systems are the teachers highly dissatisfied with the curriculum. Because of the relatively narrow range of the mean scores and because of the ambiguity of Scale 5, interpretations must be highly tentative. Still, these findings seem to offer some support for the inference made previously that a very high degree of satisfaction with the curriculum is not conducive to wholehearted effort toward curriculum change.

Summary. A comparison of the responses of teachers in two school systems indicates that the scales will differentiate schools, at least in extreme cases. The scales do not provide information about the causes of differences in the way teachers feel about the various components measured by the scales. When a number of profiles of individual schools together with case material from the schools are available for checking interpretations of new profiles, curriculum workers should be able to develop considerable skill in estimating the morale of teachers in a school in relation to curriculum development.

Six school systems were ranked by the staff of the Institute with respect to the relative success of curriculum development programs carried on in the systems. A comparison of these ranks with ranks derived from scores on the nine scales provides some evidence for the following tentative conclusions: (1) Teacher morale as measured by the scales is related to staff judgments of the success of curriculum development programs; (2) the most successful programs of curriculum development are carried on in school systems where teachers enjoy teaching and where good human relations exist; (3) the principal is a key person in a program of curriculum development; (4) teachers' feelings about money are relatively unimportant in relation to curriculum development; and (5) a high degree of satisfaction with the curriculum is not conducive to successful curriculum development.

## SECTION V

## CONCLUSIONS

This study has demonstrated that the items of the Reaction Inventory can be grouped into meaningful subscales which provide evidence of group trends, that the responses of a broad sample of teachers can provide a frame of reference for interpreting scores of groups of teachers in individual schools or school systems, and that the scales differentiate schools in relation to an index of curriculum change. The analysis provides evidence about the morale of teachers in general and supports a number of findings of previous

studies. It indicates that an index of general job satisfaction, often used as a criterion of morale in previous studies, is not a satisfactory criterion of morale and it provides a set of differential scales which is a first approximation to a satisfactory criterion.

This study recognizes the limitations of a purely statistical approach ot the appraisal of morale. Where the findings of the analysis are supported by findings of other research workers based on direct experience in actual situations, they are stated with confidence. Where findings are not supported by such evidence, they are stated tentatively, subject to validation by further study. The evidence provided by the analysis is useful in relation to other evidence available to people who may use the scales; it cannot stand by itself. Persons who use the scales should not use them mechanically but should be alert to other evidence which may be available.

The following general conclusions seem justified by the data:

1. Morale is the result of a lot of "little" things, any one of which may be focal for a particular teacher. Individual teachers experience these little things in different combinations and react to them with different intensities, but there are consistencies revealed by the responses of groups of teachers to a series of questions which justify the grouping of these little things into components.

2. No general scale is as useful as are differential scales in providing information about how morale in any situation can be improved since the components are relatively independent. This study provides a first approximation to a set of differential scales for estimating morale.

3. The components of morale which involve human relations are relatively more important than others in relation to the success of cooperative programs of curriculum improvement. The principal of a school is the key person in the development of good human relations.

4. In a particular situation, or for a particular person, any one of the components of morale may be crucial. Persons who wish to develop the high morale required for successful curriculum improvement cannot ignore any of them just because in general human relations appears to be the most important factor in most situations.

5. Not all components of morale are measured by the scales. Evidence about others, especially relations with the central administration, relations with the community, and personal adjustment of teachers, should be obtained for a full understanding of morale in any situation.

6. The scales are only approximations. They need to be improved.

## BIBLIOGRAPHY


1. American Association of School Administrators, Morale for a Free World, Twenty-Second Yearbook (Washington, D. C.: American Association of School Administrators, 1944), 461 pp.

2. Blankenship, A. B., "Methods of Measuring Industrial Morale," in G. W. Hartmann and T. Newcomb (Editors) Industrial Conflict; a Psychological Interpretation (New York: Cordon, 1939), pp. 299-312.

3. Braem, William A. F., Teacher Morale, an Investigation of Morale Factors as Judged by 150 New Jersey School Teachers. Unpublished Ed. D. project report, Teachers College, Columbia University, 1948.

4. Davies, Daniel R. "Untapped Resources of School Boards," Teachers College Record, L (April 1949), pp. 457-463.

5. Flanagan, John C., Item Analysis by Test Scoring Machine Graphic Item Counter, Proceedings of the Educational Research Forum (New York: International Business Machine Corporation, 1940), pp. 88-94.

6. Guttman, Louis, "A Basis for Scaling Qualitative Data," American Sociological Review, IX (April 1944), pp. 139-150.

7. Hoppock, Robert, Job Satisfaction (New York: Harper and Bros., 1935), 303 pp.

8. Hoppock, Robert and Spiegler, Samuel, "Job Satisfaction: Researches of 1935-1937," Occupations, XVI (April 1938), pp. 636-643.

9. Horace Mann-Lincoln Institute of School Experimentation, "Ways of Working for Curriculum Development," Teachers College Record, XLIX (February 1948), pp. 343-346.

10. Horst, Paul, and others, The Prediction of Personal Adjustment (New York: Social Science Research Council, 1941), 455 pp.

11. Hoyt, Cyril, "Test Reliability Estimated by Analysis of Variance," Psychometrika, VI (June 1941), pp. 153-160.

12. Kelley, Truman L., "The Selection of Upper and Lower Groups for the Validation of Test Items," Journal of Educational Psychology, XXX (January 1939), pp. 17-24.

13. Kolstad, Arthur, "Employee Attitudes in a Department Store," Journal of Applied Psychology, XXII (October-November 1938), pp. 470-479.

14. Likert, Rensis, "A Technique for the Measurement of Attitudes," Archives of Psychology, No. 140 (June 1932), 55 pp.

15. McNemar, Quinn, "Opinion-Attitude Methodology," Psychological Bulletin, XLIII (July 1946), pp. 289-374.

16. Mosier, C. I., "A Note on Item Analysis and the Criterion of Internal Consistency," Psychometrika, I (December 1936), pp. 275-282.

17. NEA Research Bureau, "The Teacher Looks at Personnel Administration," NEA Research Bulletin, XXIII (December 1945), pp. 96-147.

18. Raethlisberger, F. J., Management and Morale (Cambridge, Mass.: Harvard University Press, 1941), 194 pp.

19. Raethlisberger, F. J., and Dickson, W. J., Management and the Worker (Cambridge, Mass.: Harvard University Press, 1940), 615 pp.

20. Smith, Max, The Relationship Between Item Validity and Test Validity, Contributions to Education, No. 621 (New York: Bureau of Publications, Teachers College, Columbia University, 1934), 40 pp.

21. Watson, Goodwin, "The Surprising Discovery of Morale," Progressive Education, XIX (January 1942), pp. 33-41.

22. Watson, Goodwin, "Work Satisfaction," in G. W. Hartmann and T. Newcomb (Editors), Industrial Conflict; a Psychological Interpretation (New York: Cordon, 1939), pp. 114-124.

23. Wherry, Robert J. and Gaylord, Richard H., "The Concept of Test and Item Reliability in Relation to Factor Pattern," Psychometrika, VIII (December 1943), pp. 247-264.

24. Whisler, L. D., and Remmers, H. H., "A Scale for Measuring Individual and Group Morale," Journal of Psychology, IV (1937), pp. 161-165.

# PRACTICES ASSOCIATED WITH EFFECTIVE DISCIPLINE: A DESCRIPTIVE STATISTICAL STUDY OF DISCIPLINE*

SIDNEY L. CELLER
Forest Park High School
Baltimore, Maryland

## Introduction and Scope

This study took its genesis originally with the provocative teaching problem of discipline perpetually facing old and new teachers alike. It has been the sincere conviction of the author through the years of teaching and graduate study that the problem of discipline can be met best not by authoritarian control or robot bullying, but through well-planned efficient teaching by a real human being, keenly aware of the sensitivities, ambitions, and foibles of those seated before him, able to sense the humor and pathos of their lives, and willing to make their learning and personality growth his fundamental aim in teaching.

It was decided to make this an humble experimental study testing practices associated with effective discipline. It was decided to draw up a list of some thirty-eight observable practices with room for independent observations and match the frequency of their occurrence with the principal's rating in discipline of the teacher observed. This last had one chief value: it established the principal's rating as the independent criterion of discipline in each case. Thus, for the purposes of this study the concept of discipline will be considered simply as the element of classroom control a principal has in mind when he answers the questions, "Which of your teachers have the least difficulty managing their classes? Which have the most?"

Once the rating chart of some thirty-eight practices was definitely formulated, (See Appendix B)** there remained, among other things, the task of testing it in the actual classroom. To accomplish the step it was decided to observe a number of teachers, rating them on the practices listed on the chart, and then to compare these ratings of individual practices with general discipline ratings of the same teachers by their respective principals.

These ratings were to be indicated against each teacher's number, after which the principal was to destroy the key bearing each teacher's name against the appropriate number, and return the rated numbers to the observer. Strict anonymity of these teachers was preserved by the use of two sheets in each school; one with the teachers' names and their keyed numbers, the other with just the numbers and blanks for the principal's ratings. The first was destroyed by the principal in each case after the second was filled out and returned to the observer.

From the lists of volunteering teachers, eighty-two teachers of the various academic subjects were selected for observation. Sixty-six of these teachers were willing to be observed twice. All in all, then, one hundred and fifty observations of the normal forty-minute period length were made in the six schools of the eighty-two teachers participating.

## Results

### A. Objectivity

The results of an investigation of this type need to be subjected to scrutiny for their objectivity. The question is: Might the ratings be characteristic of the particular observer and not of the teaching situation nor of the teacher at all? Is what this observer sees what other observers see? In essence, how objective are the observer's ratings?

In order to anser these questions, twenty additional teachers were observed concurrently by this observer and by either one of four accredited supervisors.[1] Each supervisor used the rating chart of this study, as she sat with the ob-

---

* A summary of a doctoral dissertation submitted to the Board of University Studies of the Johns Hopkins University, Baltimore, Maryland, 1949.

**All references to Appendices may be found in original thesis, Johns Hopkins University Baltimore, Maryland.

1. Two English supervisors and one mathematics supervisor from the junior high school division of the Baltimore Secondary schools, and one English supervisor from the

server in the back of the room in the dual observations. Each supervisor chose several teachers she knew well (three by the first English supervisor, eight by the only mathematics supervisor, three by the second English supervisor, and six by the third English supervisor) and who were willing to participate, so that the observer went along on twenty dual observations. To make this perfectly clear, the maximum number of visitors to any one room in each of these twenty visits was two: the observer and one of the supervisors. It is important to note two things on these visits: (1) the supervisors were given no help with the chart except to have a copy of the Supplementary Clarification of Practices sheet, nor was the chart discussed before, during, or after a lesson—to avoid any influence from the observer in the supervisors' ratings, and (2) the supervisors did not use the chart in class but made the ratings from memory after class, sending their rating sheets on to the observer the next day without comment.

All twenty rating sheets of the supervisors were lumped together into a composite supervisors' ratings and matched with the observer's ratings for the same twenty teachers to give rise to the series of objectivity correlation coefficients presented in Column A of Table I. The data in Column B will be discussed later.

The results of this objectivity quest show striking trends. For instance, the overt teacher-techniques and personality practices (those that can be readily seen and identified) show the greatest agreement between supervisors and observer. First is Practice No. 16 "Meets disciplinary situation" with an r of .90 indicating almost perfect agreement between the supervisors and the observer as to whether teachers meet or ignore the challenge of incipient misbehavior. Since this practice refers to open measures taken against overt or obvious challenge, the extremely "high" objectivity is understandable. One can safely say that in meeting disciplinary situations the supervisors tended to see what the observer saw.

Next is Practice No. 5 "Makes use of equipment and visual aids" with an r of .88 which indicates the agreement between supervisors and observer on whether a teacher uses the facilities around her to enrich her lesson or not. The presence or absence of this practice can be noticed easily, and there is nothing esoteric about the terms "equipment" or "visual aids": hence the high agreement. There is very little, if anything, the observer would see in this connection that the supervisors would not see and vice versa.

The open use of a look or a smile to quell incipient misbehavior or the absence of this use is another practice generally agreed upon as indicated by an r of .86 for Practice No. 18 "Uses simple devices". The supervisors and the observer seem very much agreed as to when a teacher uses these simple devices to a great degree, to a fair degree, or not at all. This same high level of agreement is operative in Practice No. 20 "Uses blackboard without losing pupils" with an objectivity r of .86. The teacher's manner of using the blackboard or having the pupils use it is fairly overt and easily discernible both to the supervisors and the observer.

Practice No. 34 "Provides elastic endings" and Practice No. 35 "Gives summary" with r's of .73 and .71, respectively, are another pair of fairly openly discernible practices both to the supervisors and to the observer. A teacher can be fairly objectively observed to be giving summaries or having pupils give them, and providing elastic endings or not to be doing these things much at all.

A pair of practices which show an interesting divergence are Practice No. 12 "Avoids lulls or idleness" with an r of .72, and Practice No. 11 "Plans worthwhile work to keep each pupil challenged constantly" with an r of .54. It suggests that the supervisors and the observer are more agreed on whether a teacher makes smooth transitions from one activity to the next without lulls or idleness or not than on the apparently similar problem of whether the teacher challenges each pupil with worthwhile work or not.

The routinization of various classroom procedures and the inculcation of well-developed habits of work shows fair agreement with an r of .68 in Practice No. 7 "Develops classroom mechanics." The observance or non-observance of this practice seems fairly obvious to on-lookers, whether a teacher has her class do things smoothly in the way of drills, assignments, the handling of papers, and the use of notebooks or whether these things constitute problems that lead to confusion each time they come up.

Not as high is the agreement on Practice No. 9, "Uses subject matter appeal as a control." An objectivity r of .59 indicates only fair agreement among supervisors and observer as to whether a teacher keeps her class vitally interested in the learnings presented and impatient at disturbances or not.

A very similar condition to the above is indicated by an r of .51 for Practice No. 13 "Classroom manner free rather than restricted." The observer felt that in many cases this naturalness

Sisters of Mercy, Mount St. Agnes School, composed the group of four supervisors. In order to minimize any confusion about the terms used in the wording of each practice, each supervisor was given a sheet called "Supplementary Clarification of Practices" (See Appendix C).

TABLE I

CORRELATION COEFFICIENTS INDICATING OBJECTIVITY AND RELIABILITY FOR EACH PRACTICE

| Practice | Column A Objectivity | Column B Reliability |
|---|---|---|
| 1. Makes presence felt in pleasant way | .322 | .508 |
| 2. Begins when willingness is shown | .40 | .623 |
| 3. Takes steps to establish rapport | .072 | .755 |
| 4. Has whole class active in recall | .068 | .501 |
| 5. Makes use of equipment and visual aids | .884 | .683 |
| 6. Cares for physical conditions of room | .425 | .545 |
| 7. Develops classroom mechanics | .675 | .535 |
| * | | |
| 9. Uses subject matter appeal as control | .592 | .658 |
| 10. Uses mental hygiene values | .052 | .669 |
| 11. Plans worthwhile work | .537 | .427 |
| 12. Avoids lulls or idleness | .718 | .478 |
| 13. Classroom manner free | .51 | .57 |
| * | | |
| 15. Displays sense of humor | .369 | .384 |
| 16. Meets disciplinary situation | .90 | .464 |
| 17. Puts whole group above individual | .324 | .596 |
| 18. Uses simple devices | .865 | .668 |
| * | | |
| 20. Uses blackboard without losing pupils | .858 | .758 |
| 21. Provides for individual instruction | .373 | .682 |
| 22. Explains points carefully | .485 | .495 |
| 23. Answers points of confusion patiently | .437 | .492 |
| 24. Makes pupils feel the importance of the work | .338 | .683 |
| * | | |
| 31. Shows awareness of differences in ability | .374 | .775 |
| * | | |
| 34. Provides for elastic endings | .73 | .516 |
| 35. Gives summary or has pupils give it | .706 | .175 |

* Other practices deleted because of insufficient data; either the practice was included too late or it was not susceptible to observations.

and charm were lost once the classroom door was closed.

Next come a series of nine practices, all ranging in objectivity between .30 and .50, showing very low agreement. There are many possible explanations for this low objectivity. Perhaps the supervisors were influenced by their previous knowledge of the teachers observed. Another explanation for the low r's here may be the fact that the traits are not objectively determinable. It is perfectly possible, too, that the supervisors and the observer may have seen different things. Further, even with a supplementary sheet explaining the terms in the practices, the supervisors may have had different ideas about the meaning of these terms.[2] The overwhelming fact seems to be that these traits are not generally observed in the same way by different people. In fact, the traits as presently described may be not traits in the teacher but rather traits of which the observer is fond. These are some of the plausible hypotheses; there are others, all subject to test in future studies. For example, a trait may not exist as a characteristic trait (a questionnaire or additional observations for 'characteristicness' would help here). Also it may be that the description of a trait lacks precise landmarks whereby the trait may be identifield by the general observer. (Alternate rephrasings and submission to educators throughout the country for choice of clarity and elucidation of meaning would help here.) Finally, the possibility that previous association of the supervisors with the teachers observed may be accounting for low objectivity may be explored in another investigation by having supervisors unknown to the teachers do the rating.

Since it was not feasible for the author to have a panel of raters, this investigation is an attempt to find some traits in which the ratings of the investigator were representative, and consistently representative, of the general observer.

### B. Reliability

The matter of the consistency of a particular teacher in the practices seen by the observer comes to mind now. The question is, "Might the particular rating of a practice be characteristic of only the particular teaching situation but not consistent for this teacher?" For example, a teacher may use an excellent and unusual visual aid to make a single point, and never use another visual aid for the rest of the year.

In order to measure this repeatability of the observer's ratings, sixty-six[3] teachers were observed twice, the second visit coming not later than three weeks after the first—for the most part the next day or in the same week, teaching the same subject but not the same lesson, and not necessarily to the same grade or ability group. First observations were matched against second observations in the ratings of each practice for each of these teachers, giving rise to a series of correlation coefficients[4] presented in Column B of Table I.

It is important at the outset to note that the use of an r for reliability in this study will give rise to only a rough approximation of that reliability and a somewhat inadequate picture of it. In effect, what we are doing here is putting qualitative measures through quantitative channels, which may be convenient but is improper for intrinsic relationships. The r probably should not be used because, in the first place, we are concerned here with two qualitative scales. Secondly, the nature of the distributions are nowhere determined. Last, the distributions are in groupings of five categories which do not lend themselves adequately to precise statistical treatment.

On the other hand, the r is used here because it is convenient. Secondly, it gives some measure of reliability. Lastly and most strongly, it is used here not for precise analytical judgments, but rather for the simple presence or absence of reliability, such reliability to be used as a rough screening device for consideration or non-consideration of each practice for later treatment.

It is possible that the r's presented in Table I may have been reduced by the broad groupings used. A precise correction would tend to overcome this lack of refinement and raise the r. Assuming the distribution is normal, and referring

2. Here is a possible explanation for low objectivity borne out by the data: the variation from supervisor to supervisor. It may be the fact that different teachers were observed by different supervisors. Or it may be the possibility mentioned above about different ideas. Consult Appendix D for actual ratings.

3. The other sixteen teachers (to make a total of eighty-two) either would not consent to a second observation or could not be reached because of time or schedule limitations.

4. The Pearson product-moment formula was used. See C. C. Peters and W. R. VanVoorhis, Statistical Procedures and Their Mathematical Bases (New York: McGraw Hill, 1940), p. 100.

to Peters and Van Voorhis' table[5] for the r between index values and variates in case of five categories, the r of .18 for practice number 35 would become .20, a correction amounting to no more than two or three points.

Another consideration is the possibility that the standard error of the reliability r might prove to be too large. The question is whether, for each practice, we could reasonably expect to have obtained as large an r as we did if the true reliability r is zero. Using the formula[6]

$$\sigma r = \frac{1}{\sqrt{N - 1}}$$

the standard error of the r, in our case, where the true r is zero, comes out to be .12. In light of this small standard error, and the fact that the lowest reliability used for significant results was .46, we can say that our obtained r differs reliably from zero.

A discussion of the possibilities pregnant in either a low or high reliability is in point here. Low reliability could be due to, first, an actual change in the status or behavior of the teacher from one visit to the next, this change being accurately reflected by the rater. Secondly, it is possible that the actual trait itself was fairly constant but that the rater observing this trait put down the rating inaccurately. Or it might be the rater was really accurately observing some extraneous trait which fluctuated. Or it might be that the rater was observing a fictional trait, wherein no real trait could be observed. Because of the subjective nature of the rating situation, the most plausible of the above hypotheses lie in the inaccuracy of the rater in one way or another.

The possible explanations for whatever high reliability is found are just as numerous. It might be that the observer was rating the actual trait underlying the description, and accurately rating it, with the trait remaining fairly constant. Or it might be that the rater was again observing the actual trait, which itself was not constant, but the apparent constancy shown in the rating coming about because the rater was depending on some extrinsic trait which happened to remain constant or some subjective cues which happened to remain constant. Or it might be that the rater was rating a fictional trait with apparent constancy due to the constant subjective cues. Or it might be that the rater was simply remembering the first rating, this memory effect operating even though the observer sought conscientiously to avoid it. It is important therefore, to keep in mind that a high reliability figure is indicative only of the fact that similar ratings were given by the rater; the rating being consistent without the teacher necessarily being consistent at all.

An inspection of Column B of Table I reveals first, that some of the teacher-technique practices had higher reliability than others. These practices with higher reliability are Practice No. 20 "Uses blackboard without losing pupils" (.76), Practice No. 5 "Makes use of equipment and visual aids" (.68), and Practice No. 21 "Provides for individual instruction" (.68). In this same group are to be found practices representing teacher-personality traits. They are Practice No. 31 "Shows awareness of differences in ability" (.78), Practice No. 3 "Takes steps to establish rapport" (.76), Practice No. 24 "Makes pupils feel the importance of the work" (.68), Practice No. 10 "Uses mental hygiene values for each pupil" (.67), Practice No. 18 "Uses simple devices, a look or a smile" (.67), and Practice No. 9 "Uses subject matter appeal as a control" (.66).

Although the coefficients presented above are not very high, they refer to practices which were rated most similarly by the observer in the series of sixty-six double observations. To account for this similarity of rating, the hypotheses advanced for high reliability may be applied. For example, in explaining an r of .78 for Practice No. 31 "Shows awareness of differences of ability," it may be that the observer was rating the actual awareness the teachers showed to different abilities in their groups, and rating it accurately, with the awareness remaining fairly constant. Another possibility is that the rater was observing the actual awareness, which itself was not constant, but the apparent constancy shown in the rating came about because of a memory effect, or the fact that the rater was depending upon the teachers' general manner of being concerned about all the pupils in her class which carried over from period to period. It might be, and this is most likely in light of the low objectivity r of .34, that the awareness did not enter into the rating at all, the rater basing his constant rating on the constant presence or absence of small kindnesses to the slow pupils in the class, or some other subjective cue.

Another example will suffice. An r of .76 for Practice No. 3 "Takes steps to establish rapport" is explainable in several different ways. It is possible, for instance, that the rater was observing real rapport-establishing and rating it accurately, with the rapport being fairly constant. Yet

5. Ibid., p. 398.

6. Ibid., p. 153.

this is the most improbable in light of the complete lack of objectivity for the practice (.07). The most probable explanation is that the observer was rating the presence or absence of a fictional rapport known only to him and recognized by only him through a series of personal landmarks, such as the warm smile, the constant use of "we", etc.

The only practices about which we can say with a high degree of probability that the observer was rating the actual trait accurately, with the trait remaining fairly constant, are the overt teacher-technique practices, such as Practice No. 20 "Uses blackboard" and Practice No. 5 "Makes use of equipment and visual aids" which are fairly objective (.86 and .88, respectively).

The low reliability r's are found in such practices as Practice No. 35 "Gives summary or has pupils give it" (.18), Practice No. 15 "Displays sense of humor for relevant fun" (.38), Practice No. 11 "Plans worthwhile work to keep each pupil challenged constantly" (.43), Practice No. 16 "Meets disciplinary situation" (.46), Practice No. 12 "Avoids lulls or idleness" (.48), Practice No. 22 "Explains points carefully" (.50), and Practice No. 23 "Answers points of confusion patiently" (.49).

Here the hypotheses advanced for low reliability may be applied, again with different degrees of plausibility. For example, in a practice as overt and readily recognizable (objectivity r of .71) as No. 35 "Gives summary or has pupils give it," the most plausible assumption to account for the lack of reliability is the probability that the teacher varied in her practice of giving summaries from one visit to the next, this change being accurately reflected by the rater.

However, such an explanation to account for the low reliability in Practice No. 15 "Displays a sense of humor for relevant fun" would be stretching the ground of plausibility. The low reliability for a practice as difficult of definitive recognition as this is (objectivity r of .38) is more logically explained by the probability that the rater was not observing or rating a sense of humor at all, but rather some extraneous traits like a pleasant disposition or ready smile, which in themselves fluctuate. In fact, the rater may have been observing these extraneous traits inaccurately with the traits, themselves, remaining constant. It is also possible that the ratings were based on the observer's special and personal concept of a sense of humor, dependent upon a good number of subjective cues which vary themselves.

This last possibility may account for the low reliability obtained for Practice No. 22 "Explains points carefully" and Practice No. 23 "Answers points of confusion patiently;" for it is perfectly plausible to believe that the rater may have used his personal and subjective concepts of care and patience to rate these practices. Objectivity coefficients of .49 and .44, respectively, bear out this possibility.

A final explanation for low reliability lies in the simple inaccuracy of the observer in putting down ratings. Such a possibility might well account for the low reliability obtained for Practice No. 12 "Avoids lulls or idleness." The avoiding of lulls may have been constant and generally observable (objectivity r of .72) but the observer may not have seen it, or, if he did, rated it inaccurately.

C. Relationship

The section on objectivity matched the observer's ratings in the various practices for twenty addtional teachers with a composite of supervisors' ratings for the same twenty teachers observed concurrently by the observer and one of the supervisors to establish the relative objectivity of these practices. The section on reliability matched the observer's ratings in the various practices for each of sixty-six teachers on the first observation with the observer's ratings in the same practices for the same teachers on the second observation to establish the relative reliability of these practices. Once the objectivity and reliability for each practice has been fairly well-established, the decks are cleared for the real problem of this study: what are the relationships between the principals' ratings of the teachers in discipline and the observer's ratings of these same teachers in the disciplinary practices used in this study? In essence, which traits or practices go along with high discipline ratings from the principals, and which traits or practices are typical of those teachers given low discipline ratings by the principals?

In the consideration of relationship in the immediately succeeding pages, objectivity and reliability are temporarily somewhat neglected until a later synthesis. This is done because even when objectivity and reliability are low, a high relation has some theoretical significance and should be suggestive or act as a lead. Furthermore, standards of acceptable objectivity and reliability may vary from reader to reader. With a complete presentation every man may select his own standards.

The data were attacked in three different ways, all related and all diffusing some light on this problem of the meaningful relationship between principals' assessment of their teachers' disciplinary abilities and the disciplinary practices used or neglected by these teachers as seen by the observer. The most direct attack is simple inspection of the correlation charts showing the observer's ratings in each of the twenty-four

practices[7] (see Tables II through XXV[8]) in relation to the principals' ratings. In these charts the x-axis represents variations in the different disciplinary ability as rated by the principals. Thus, one can see how well these observer ratings in each practice parallel the conception of disciplinary ability the principals have for these teachers. For example, by inspection, one can tell how the teachers rated E in disciplinary ability by the principals have been rated by the observer in any particular practice. This can be done for any rating and in any practice by inspection of the frequency tables. Only the first observations of all eighty-two teachers (exclusive of the twenty used in gathering the objectivity data) involved in the study were used. This restriction was employed to make sure that no double weighting would occur because of repetitive ratings. It is also important to note that the eighty-two teachers were rated in the disciplinary ability by the principals as follows: Poor 13, Fair 13, Good 22, Very Good 5, and Excellent 29 (these are the P, F, G, VG, and E headings in each table to make a total N of 82), making for three ability groups of teachers (P plus F 26, G plus VG 27, and E 29) of equal size. The left-hand vertical column lists the possible observer's ratings; and the last horizontal column, labeled M, gives the weighted averages of these observer's ratings among the various principals' ratings divisions. By inspection of these averages reading from left to right one can tell whether the observer's ratings in any one of the 24 practices increased with the increased principals' estimates of the disciplinary abilities of the teachers or not.

These averages were used to plot a series of straight-line graphs so that the relationship could be seen more clearly. This second method of attack is presented in Figures 2 through 25, facing the frequency distributions from which the graphs derive their values.[9] The principals' ratings are laid out along the abscissa, and the observer's ratings make up the ordinate intervals. The weighted averages of the observer's ratings are plotted at the five principals' rating points and connected by a continuous line to show the slope or direction of these values. To discount the disproportionate effect of the VG category which contains only five cases and to eliminate other minor irregularities, the data were then smoothed into three large groups mentioned above, plotted on the same graph, and connected by a broken line. In general, the broken line will accentuate the steepness of the continuous line slope or the lack of steepness. The reader will find the closest relationship between the principals' and observer's ratings where this slope is steepest, because a steep slope indicates, for the most part, that the teachers rated P by the principals were rated 1 by the observer, the teachers rated F by the principal were rated 2 by the observer, etc.

The third and final method of attack on this problem of relationship was the use of the chi-square[10] test on each practice. This statistical device was used here to test the hypothesis that the variates in each of the contingency tables were significantly related. It must be remembered that this is a general technique of analysis, yielding no direction to the relationship. A high chi-square, for example, might be indicative of a close but inverse relationship. It does, however, provide a probability figure (P) as to pure chance implications in each case.

Since the theoretical frequency is the point of reference in the calculation of chi-square, it is important that this theoretical frequency for each cell be large enough to provide stability. Peters and VanVoorhis imply a theoretical frequency of 10 as a comfortable minimum and 5 as an absolute minimum.[11] Because of an unfortunate oversight this condition was not met in the original calculation of chi-square. Instead, the computations were based on the untreated tables.[12] When the oversight was detected, chi-square was recomputed. In this recomputation the groupings were made coarse enough to insure a minimum of 5 for each theoretical frequency. This, of course, resulted in fewer degrees of freedom. The new

7. The practices are presented in the rank order of their objectivity coefficients assuming a fairly high reliability coefficient.

8. For ready location of the figure and table for each practice, see accompanying auxiliary key.

9. Figure 2 is the first figure even though there is no figure 1 in the study. This was done for convenience of matched numbers since the corresponding frequency distribution had to be labeled Table II.

10. Peters and VanVoorhis, op. cit., Ch. XIV.

11. Peters and VanVoorhis, op. cit., p. 417.

12. See Tables II - XXV.

## AUXILIARY KEY TO FIGURES AND TABLES FOR EACH PRACTICE

| Practice No. | Figure | Page | Table | Page |
|---|---|---|---|---|
| 1 | 20 | 348 | 20 | 349 |
| 2 | 15 | 346 | 15 | 347 |
| 3 | 23 | 348 | 23 | 349 |
| 4 | 24 | 348 | 24 | 349 |
| 5 | 2 | 342 | 2 | 343 |
| 6 | 14 | 346 | 14 | 347 |
| 7 | 5 | 342 | 5 | 343 |
| 9 | 7 | 342 | 7 | 343 |
| 10 | 25 | 348 | 25 | 349 |
| 11 | 10 | 344 | 10 | 345 |
| 12 | 9 | 344 | 9 | 345 |
| 13 | 11 | 344 | 11 | 345 |
| 15 | 21 | 348 | 21 | 349 |
| 16 | 8 | 344 | 8 | 345 |
| 17 | 19 | 346 | 19 | 347 |
| 18 | 3 | 342 | 3 | 343 |
| 20 | 4 | 342 | 4 | 343 |
| 21 | 17 | 346 | 17 | 347 |
| 22 | 12 | 344 | 12 | 345 |
| 23 | 13 | 344 | 13 | 345 |
| 24 | 18 | 346 | 18 | 347 |
| 31 | 16 | 346 | 16 | 347 |
| 34 | 6 | 342 | 6 | 343 |
| 35 | 22 | 348 | 22 | 349 |

chi-squares and the corresponding P values appear in Appendix E (see original thesis).

It happened that some of the 24 P values were changed as a result of the recomputation. In the tables and throughout the discussion these changed values are indicated by *-marked footnotes.[13]

Using this three-way analysis on the actual results, the reader's attention is called first to the frequency distribution of Practice No. 5 "Makes use of equipment and visual aids" given in Table II. Of the 14 teachers rated 5 by the observer in the use of this practice, 10 were E teachers in discipline as far as the principals were concerned. Nine of the 16 rated 4 by the observer were E teachers. Five of the 10 teachers given a 1 by the observer were rated either P or F by the principals. The relationship seems fairly close. The graph in Figure 2 emphasizes this agreement only mildly because the rise in the slope is not very steep, going up only 1.2 units from 2.5 to 3.7, attributable perhaps to the two P teachers rated 5 by the observer and the 10 E teachers rated 3 or below. However, the chi-square further corroborates the relationship noticed at first by yielding a high figure of 44.5 and a P low enough to indicate that there are only four chances in 10,000 that a chi-square as large as this could have occurred by chance and that there is a genuine relationship between the two variables: the principals' ratings and the observer's ratings. Summing up, making use of equipment and visual aids seems to be one practice definitely associated with the principals' concept of effective discipline, at least as far as these teachers are concerned.

Practice No. 18 "Uses simple devices" shows a relationship between the two variables but a less definite one. The frequency distribution for this practice, reproduced in Table III, does show that out of the 13 teachers rated 5 by the observer, 11 were E teachers; but only 6 of the 14 teachers rated 4 were E teachers. In addition while 19 of the 30 teachers rated 1 were P or F calibre according to the principals, 6 of the 30 were of E disciplinary ability. Thus the relationship is not as close as the one depicted above. The graph in Figure 3 gives a more favorable picture of the relationship, revealing a steeper slope than in Figure 2 primarily because 10 of the 13 P teachers were rated 1 to make a low weighted average at the beginning of the graph, a fairly even distribution of the 13 G teachers in all five observer ratings to make for a fairly steady rise, and a concentration of 17 of the 29 E teachers in ratings 5 and 4 to make a high weighted average at the end of the graph. The line rises 1.8 units from start to finish and reveals a closer relationship than the frequency distribution indicates. The chi-square yields a middle figure of 26.7 with a P high enough to indicate that there are as many as 4-1/2* chances in 100 that this chi-square could have occurred by chance. All together, we can say that such a probability gives us only moderate confidence that the relation of the practice of using simple devices such as a look or a smile with our principals' concept of effective discipline is not due to chance.

Practice No. 20 "Uses blackboard without losing pupils" gives the same type of picture, with even less of a relationship indicated. The frequency distribution in Table IV shows that 8 of the 12 teachers rated 5 were E teachers, but those rated 4 are spread fairly evenly throughout all the different abilities and the P, F, and G teachers seem to share equally those teachers rated 1 by the observer. The concentration of 10 of the 14 E teachers in ratings 5 and 4 making for a high weighted average at the end, is responsible for the surprising steepness of the slope and the rise of 2.1 units in Figure 4. The chi-square of 22.5 gives perhaps the truest picture with the probability that there are as many as 13* chances in 100 that the relationship between the observer's ratings and the principals' ratings in this practice could have occurred by chance. In view of this probability, the relationship must be regarded as having questionable significance.

Practice No. 7 "Develops classroom mechanics" reveals a much tighter and closely related picture. Table V indicates that as many as 15 of the 26 teachers rated 5 were E teachers, 9 of the 15 teachers rated 4 were of VG or E calibre, and 14 of the 23 rated 1 were either P or F teachers. The graph in Figure 5 does not reveal as steep a slope as might be expected largely because there are 4 cases of P teachers rated 5 which make for a rather high weighted average at the beginning. The smoother line irons out this weighting and gives a clearer picture of the 1.6 units of rise involved in the slope. The chi-square test indicates the fairly high figure of 34.2 with the probability that there is only one* chance in 100 that a chi-square as large as this could have occurred by chance. Here again, as in the first practice discussed, one can say that the observer's rating in the practice of developing classroom mechanics seems definitely associated with effective discipline as defined by the principals' ratings.

We go down again slightly in the degree of relationship in the examination of Table VI, the

13. See Appendix E in original thesis for a full table of these and the original values.

* Accurate value of P is less than .01.

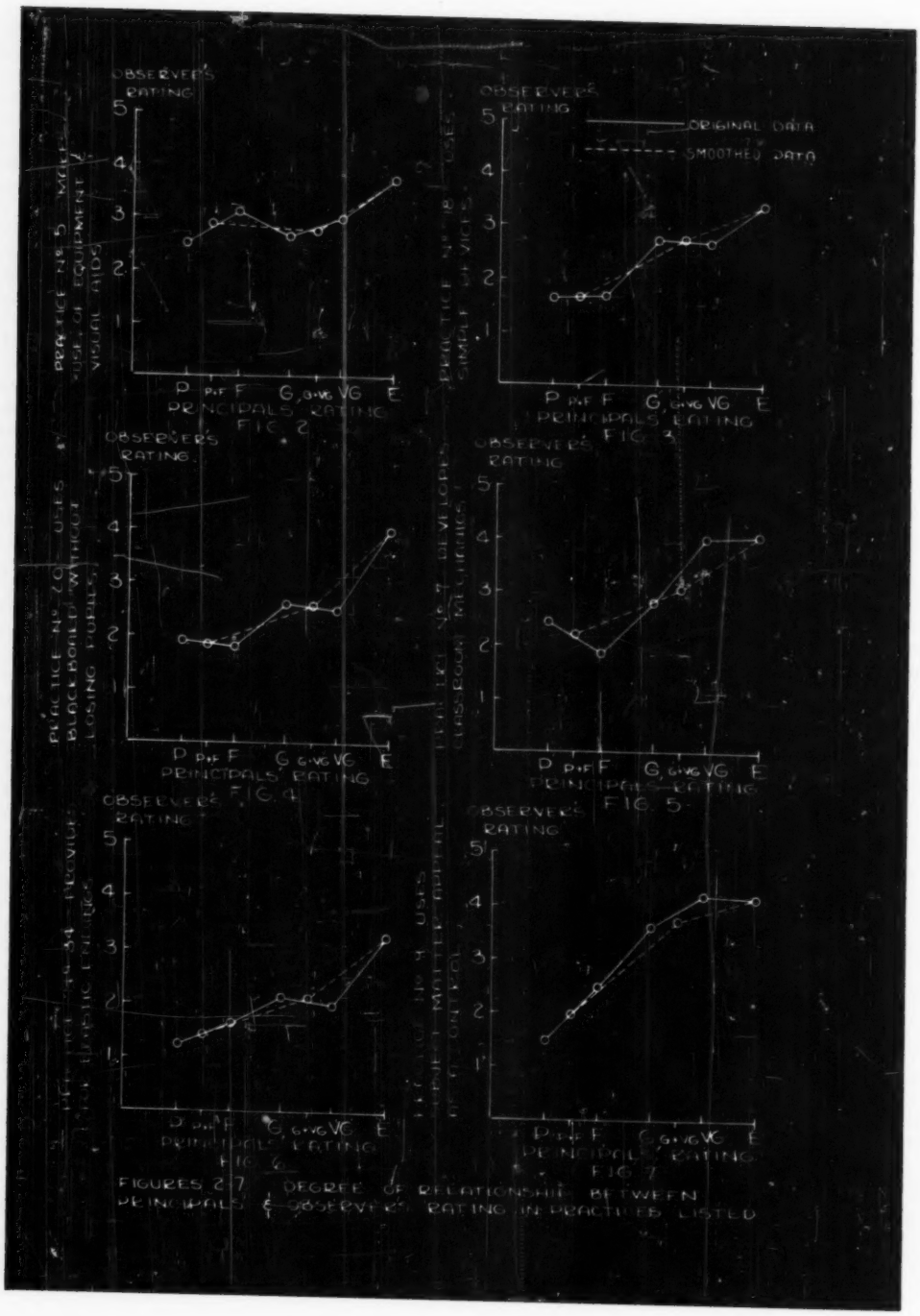

OBSERVER'S RATING
PRACTICE NO. 5 MAKES USE OF EQUIPMENT VISUAL AIDS
P P+F F G G+VG VG E
PRINCIPALS' RATING
FIG. 2
ORIGINAL DATA
SMOOTHED DATA
PRACTICE NO. 8 USES SMALL GROUPS
FIG. 3
PRACTICE NO. 20 USES BLACKBOARD WITHOUT LOSING PUPILS
FIG. 4
PRACTICE NO. 7 DEVELOPS CLASSROOM MECHANICS
FIG. 5
PRACTICE NO. 34 PROVIDES ESSENTIAL FACILITIES
FIG. 6
PRACTICE NO. 4 USES SUBJECT MATTER APPEAL
FIG. 7
FIGURES 2-7 DEGREE OF RELATIONSHIP BETWEEN PRINCIPALS' & OBSERVERS' RATING IN PRACTICES LISTED

TABLE II[13a]

FREQUENCY DISTRIBUTION OF PRACTICE NO. 5: "EQUIPMENT AND VISUAL AIDS"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 2 | 2 | 0 | 0 | 10 | 14 |
| 4 | 0 | 3 | 2 | 2 | 9 | 16 |
| 3 | 1 | 0 | 7 | 2 | 4 | 14 |
| 2 | 2 | 2 | 3 | 0 | 4 | 11 |
| 1 | 3 | 2 | 2 | 1 | 2 | 10 |
| N | 8 | 9 | 14 | 5 | 29 | 65* |
| M | 2.5 | 3.1 | 2.6 | 3.0 | 3.7 | |

Chi-square = 44.5; P = less than .01

TABLE III

FREQUENCY DISTRIBUTION OF PRACTICE NO. 18: "USES SIMPLE DEVICES"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 0 | 2 | 0 | 11 | 13 |
| 4 | 2 | 2 | 3 | 1 | 6 | 14 |
| 3 | 1 | 1 | 2 | 1 | 2 | 7 |
| 2 | 0 | 1 | 2 | 0 | 4 | 7 |
| 1 | 10 | 9 | 4 | 1 | 6 | 30 |
| N | 13 | 13 | 13 | 3 | 29 | 71* |
| M | 1.6 | 1.7 | 2.8 | 2.7 | 3.4 | |

Chi-square = 26.7; P = between .05 and .01**

TABLE IV

FREQUENCY DISTRIBUTION OF PRACTICE NO. 20: "USES BLACKBOARD NOT LOSING PUPILS"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 1 | 0 | 3 | 0 | 8 | 12 |
| 4 | 1 | 2 | 2 | 2 | 2 | 9 |
| 3 | 1 | 1 | 1 | 0 | 1 | 4 |
| 2 | 1 | 0 | 1 | 0 | 2 | 4 |
| 1 | 7 | 7 | 6 | 2 | 1 | 23 |
| N | 11 | 10 | 13 | 4 | 14 | 52* |
| M | 1.9 | 1.8 | 2.6 | 2.5 | 4.0 | |

Chi-square = 22.5; P = greater than .10**

TABLE V

FREQUENCY DISTRIBUTION OF PRACTICE NO. 7: "DEVELOPS CLASSROOM MECHANICS"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 4 | 1 | 4 | 2 | 15 | 26 |
| 4 | 1 | 1 | 4 | 2 | 7 | 15 |
| 3 | 0 | 0 | 5 | 0 | 2 | 7 |
| 2 | 1 | 4 | 2 | 1 | 3 | 11 |
| 1 | 7 | 7 | 7 | 0 | 2 | 23 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 2.4 | 1.8 | 2.8 | 4.0 | 4.0 | |

Chi-square = 34.2; P = between .05 and .01**

TABLE VI

FREQUENCY DISTRIBUTION OF PRACTICE NO. 34: "PROVIDES FOR ELASTIC ENDINGS"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 0 | 4 | 1 | 14 | 19 |
| 4 | 1 | 1 | 1 | 0 | 3 | 6 |
| 3 | 0 | 2 | 2 | 0 | 0 | 4 |
| 2 | 0 | 1 | 2 | 1 | 1 | 5 |
| 1 | 12 | 9 | 13 | 3 | 11 | 48 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.2 | 1.6 | 2.1 | 2.0 | 3.3 | |

Chi-square = 30.3; P = between .05 and .01**

TABLE VII

FREQUENCY DISTRIBUTION OF PRACTICE NO. 9: "USES SUBJECT MATTER APPEAL"

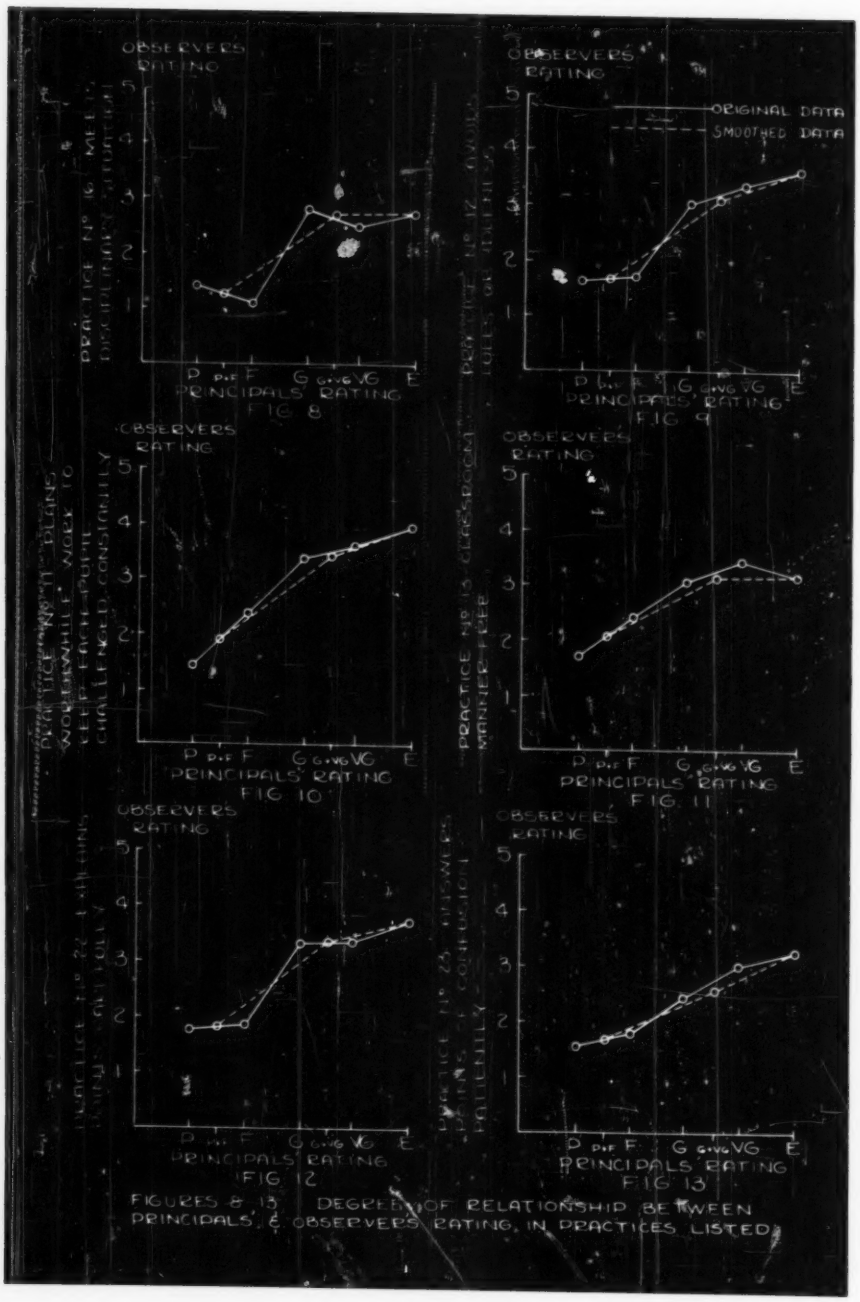
| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 1 | 2 | 10 | 3 | 17 | 33 |
| 4 | 0 | 2 | 4 | 1 | 5 | 12 |
| 3 | 0 | 1 | 2 | 0 | 3 | 6 |
| 2 | 2 | 3 | 2 | 1 | 2 | 10 |
| 1 | 10 | 5 | 4 | 0 | 2 | 21 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.5 | 2,5 | 3,6 | 4.2 | 4.1 | |

Chi-square = 36.4; P = less than .01

[13a] For ready location of figures and tables for each practice, see accompanying auxiliary key on page 340

* Other cases not susceptible to observation.

**For accurate value of P based upon coarser groupings, see Appendix E in original thesis, Johns Hopkins University.




FIGURES 8 TO 13 DEGREE OF RELATIONSHIP BETWEEN PRINCIPALS' & OBSERVERS' RATING, IN PRACTICES LISTED

TABLE VIII

FREQUENCY DISTRIBUTION OF PRACTICE NO. 16: "MEETS DISCIPLINARY SITUATION"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 0 | 2 | 0 | 3 | 5 |
| 4 | 1 | 0 | 1 | 0 | 3 | 5 |
| 3 | 0 | 0 | 4 | 1 | 2 | 7 |
| 2 | 1 | 1 | 2 | 1 | 1 | 6 |
| 1 | 8 | 9 | 3 | 0 | 6 | 26 |
| N | 10 | 10 | 12 | 2 | 15 | 49* |
| M | 1.4 | 1.1 | 2.8 | 2.5 | 2.7 | |

Chi-square = 25.6; P = between .10 and .05***

TABLE IX

FREQUENCY DISTRIBUTION OF PRACTICE NO. 12: "AVOIDS LULLS OR IDLENESS"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 2 | 1 | 5 | 1 | 13 | 22 |
| 4 | 0 | 1 | 6 | 2 | 6 | 15 |
| 3 | 0 | 0 | 3 | 1 | 2 | 6 |
| 2 | 0 | 2 | 1 | 0 | 2 | 5 |
| 1 | 11 | 9 | 7 | 1 | 6 | 34 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.6 | 1.7 | 3.0 | 3.4 | 3.6 | |

Chi-square = 31.2; P = between .05 and .01***

TABLE X

FREQUENCY DISTRIBUTION OF PRACTICE NO. 11: "PLANS WORTHWHILE WORK"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 2 | 6 | 2 | 15 | 25 |
| 4 | 1 | 1 | 8 | 1 | 4 | 15 |
| 3 | 1 | 2 | 1 | 1 | 6 | 11 |
| 2 | 0 | 3 | 2 | 0 | 1 | 6 |
| 1 | 11 | 5 | 5 | 1 | 3 | 25 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.4 | 2.4 | 3.4 | 3.6 | 3.9 | |

Chi-square = 41.4; P = less than .01

TABLE XI

FREQUENCY DISTRIBUTION OF PRACTICE NO. 13: "CLASSROOM MANNER FREE"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 1 | 2 | 0 | 6 | 9 |
| 4 | 2 | 4 | 9 | 3 | 8 | 26 |
| 3 | 0 | 1 | 3 | 1 | 6 | 11 |
| 2 | 3 | 0 | 4 | 1 | 2 | 10 |
| 1 | 8 | 7 | 4 | 0 | 7 | 26 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.7 | 2.4 | 3.0 | 3.4 | 3.1 | |

Chi-square = 34.4; P = between .10 and .05***

TABLE XII

FREQUENCY DISTRIBUTION OF PRACTICE NO. 22: "EXPLAINS POINTS CAREFULLY"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 0 | 0 | 1 | 4 | 5 |
| 4 | 1 | 0 | 1 | 1 | 1 | 4 |
| 3 | 0 | 1 | 2 | 0 | 0 | 3 |
| 2 | 0 | 3 | 0 | 0 | 0 | 3 |
| 1 | 3 | 2 | 0 | 1 | 2 | 9 |
| N | 4 | 6 | 3 | 3 | 7 | 23** |
| M | 1.8 | 1.8 | 3.3 | 3.3 | 3.7 | |

Chi-square = 28.8; P = between .05 and .01

TABLE XIII

FREQUENCY DISTRIBUTION OF PRACTICE NO. 23: "ANSWERS CONFUSION PATIENTLY"

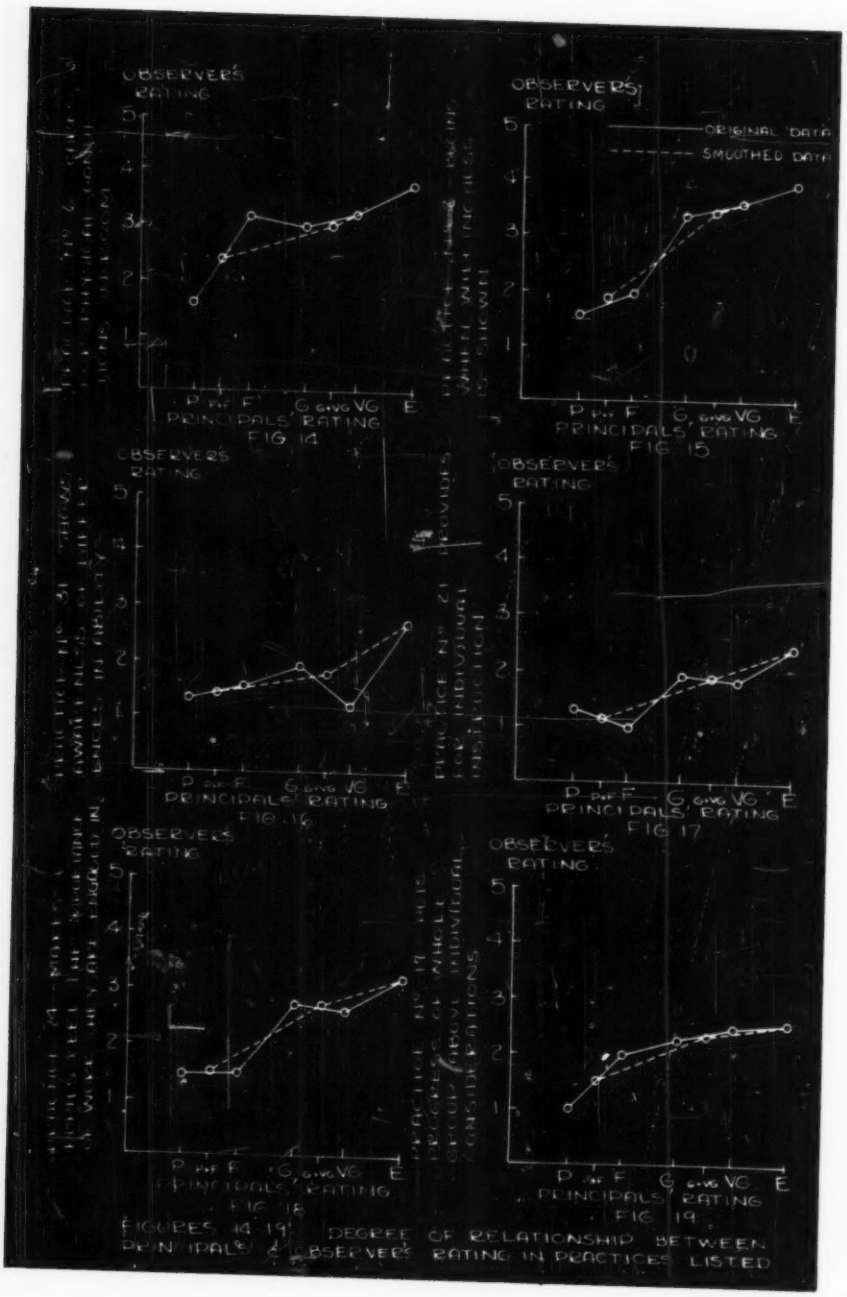
| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 1 | 1 | 3 | 1 | 13 | 19 |
| 4 | 1 | 2 | 5 | 1 | 3 | 12 |
| 3 | 0 | 0 | 1 | 1 | 2 | 4 |
| 2 | 0 | 0 | 3 | 1 | 0 | 4 |
| 1 | 11 | 10 | 1 | 1 | 11 | 43 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.5 | 1.8 | 2.4 | 3.0 | 3.2 | |

Chi-square = 31.7; P = between .05 and .01***

* Other cases not susceptible to observation.

** Practice on new form only.

***For accurate value of P based upon coarser groupings, see Appendix E in original thesis Johns Hopkins University.




FIGURES 14-19. DEGREE OF RELATIONSHIP BETWEEN PRINCIPALS' AND OBSERVERS' RATING IN PRACTICES LISTED

TABLE XIV

FREQUENCY DISTRIBUTION OF PRACTICE NO. 6: "PHYSICAL CONDITIONS OF ROOM"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 3 | 3 | 0 | 12 | 18 |
| 4 | 2 | 4 | 6 | 2 | 7 | 21 |
| 3 | 1 | 0 | 4 | 2 | 4 | 11 |
| 2 | 0 | 4 | 5 | 1 | 3 | 13 |
| 1 | 10 | 2 | 4 | 0 | 3 | 19 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.6 | 3.2 | 3.0 | 3.2 | 3.8 | |

Chi-square = 40.7; P = less than .01

TABLE XV

FREQUENCY DISTRIBUTION OF PRACTICE NO. 2: "BEGINS WHEN WILLING"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 1 | 0 | 4 | 0 | 14 | 19 |
| 2 | 1 | 1 | 9 | 4 | 6 | 21 |
| 3 | 0 | 3 | 4 | 0 | 4 | 11 |
| 2 | 0 | 3 | 1 | 1 | 3 | 8 |
| 1 | 11 | 6 | 4 | 0 | 2 | 23 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.5 | 1.9 | 3.4 | 3.6 | 3.9 | |

Chi-square = 55.5; P = less than .01

TABLE XVI

FREQUENCY DISTRIBUTION OF PRACTICE NO. 31: "SHOWS AWARENESS OF DIFFERENCES"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 1 | 0 | 3 | 0 | 9 | 13 |
| 4 | 0 | 0 | 2 | 0 | 3 | 5 |
| 3 | 0 | 1 | 1 | 0 | 1 | 3 |
| 2 | 0 | 5 | 0 | 1 | 2 | 8 |
| 1 | 12 | 7 | 16 | 4 | 14 | 53 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.3 | 1.5 | 1.9 | 1.2 | 2.7 | |

Chi-square = 30.2; P = between .05 and .01**

TABLE XVII

FREQUENCY DISTRIBUTION OF PRACTICE NO. 21: "INDIVIDUAL INSTRUCTION"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 1 | 0 | 1 | 0 | 7 | 9 |
| 4 | 0 | 0 | 4 | 0 | 3 | 7 |
| 3 | 0 | 0 | 2 | 1 | 2 | 5 |
| 2 | 0 | 0 | 0 | 2 | 1 | 3 |
| 1 | 12 | 13 | 15 | 2 | 16 | 58 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.3 | 1.0 | 1.9 | 1.8 | 2.4 | |

Chi-square = 37.8; P = less than .01

TABLE XVIII

FREQUENCY DISTRIBUTION OF PRACTICE NO. 24: "MAKES PUPILS FEEL THE IMPORTANCE OF WORK THEY ARE ENGAGED IN"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 1 | 0 | 6 | 2 | 11 | 20 |
| 4 | 1 | 2 | 3 | 0 | 4 | 10 |
| 3 | 0 | 0 | 1 | 0 | 3 | 4 |
| 2 | 0 | 0 | 3 | 0 | 2 | 5 |
| 1 | 11 | 11 | 9 | 3 | 9 | 43 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.5 | 1.5 | 2.7 | 2.6 | 3.2 | |

Chi-square = 24.6; P = between .10 and .05**

TABLE XIX

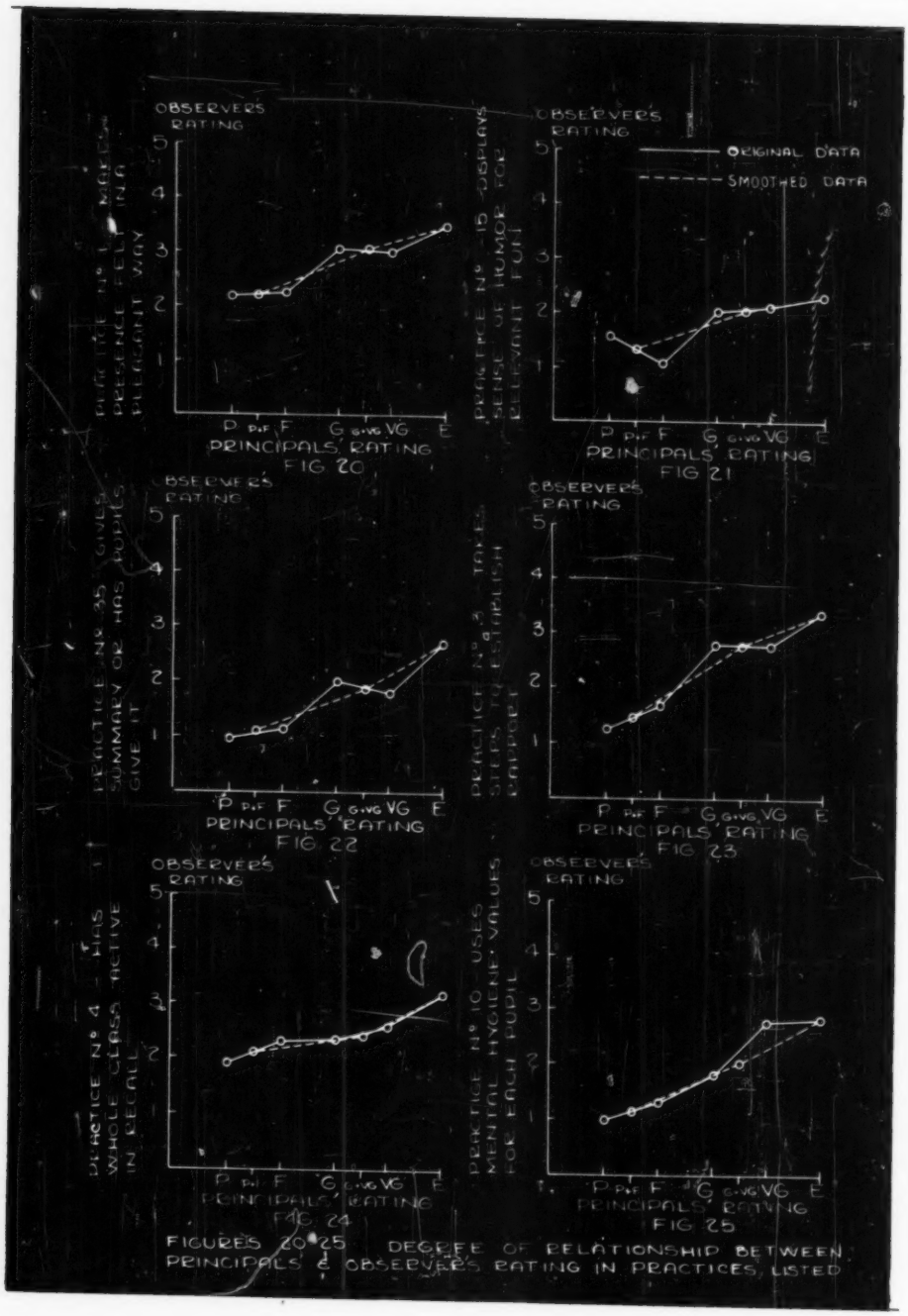
FREQUENCY DISTRIBUTION OF PRACTICE NO. 17: "PUTS PROGRESS OF WHOLE GROUP ABOVE INDIVIDUAL CONSIDERATIONS

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 0 | 0 | 0 | 2 | 2 |
| 4 | 0 | 1 | 2 | 2 | 1 | 6 |
| 3 | 0 | 1 | 3 | 0 | 4 | 8 |
| 2 | 0 | 1 | 1 | 0 | 0 | 2 |
| 1 | 6 | 6 | 4 | 2 | 5 | 20 |
| N | 6 | 6 | 10 | 4 | 12 | 38* |
| M | 1.0 | 2.0 | 2.3 | 2.5 | 2.6 | |

Chi-square = 19.0; P = greater than .10

* Other cases not susceptible to observation.

**For accurate value of P based upon coarser groupings, see Appendix E in original thesis Johns Hopkins University.


OBSERVERS RATING
PRACTICE Nº 1 MAKES PRESENCE FELT IN A PLEASANT WAY
PRACTICE Nº 15 DISPLAYS SENSE OF HUMOR FOR RELEVANT FUN
ORIGINAL DATA
SMOOTHED DATA
P Pnf F G G+VG VG E
PRINCIPALS' RATING
FIG 20
FIG 21
PRACTICE Nº 35 GIVES SUMMARY OR HAS PUPILS GIVE IT
PRACTICE Nº 3 TAKES STEPS TO ESTABLISH RAPPORT
FIG 22
FIG 23
PRACTICE Nº 4 HAS WHOLE CLASS ACTIVE IN RECALL
PRACTICE Nº 10 USES MENTAL HYGIENE VALUES FOR EACH PUPIL
FIG 24
FIG 25

FIGURES 20-25 DEGREE OF RELATIONSHIP BETWEEN PRINCIPALS' & OBSERVERS' RATING IN PRACTICES LISTED

TABLE XX

FREQUENCY DISTRIBUTION OF PRACTICE NO. 1: "MAKES PRESENCE PLEASANT"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 1 | 1 | 1 | 9 | 12 |
| 4 | 5 | 1 | 10 | 1 | 8 | 25 |
| 3 | 0 | 3 | 3 | 0 | 4 | 10 |
| 2 | 0 | 3 | 5 | 3 | 4 | 15 |
| 1 | 8 | 5 | 3 | 0 | 4 | 20 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 2.2 | 2.2 | 3.0 | 3.0 | 3.5 | |

Chi-square = 37.2; P = less than .01

TABLE XXI

FREQUENCY DISTRIBUTION OF PRACTICE NO. 15: "DISPLAYS SENSE OF HUMOR"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 1 | 0 | 2 | 1 | 4 | 8 |
| 4 | 1 | 0 | 4 | 0 | 4 | 9 |
| 3 | 0 | 0 | 1 | 1 | 4 | 6 |
| 2 | 0 | 1 | 0 | 0 | 2 | 3 |
| 1 | 11 | 12 | 15 | 3 | 15 | 56 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.5 | 1.1 | 2.0 | 2.1 | 2.3 | |

Chi-square = 15.9; P = greater than .10**

TABLE XXII

FREQUENCY DISTRIBUTION OF PRACTICE NO. 35: "GIVES SUMMARY OR HAS IT GIVEN"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 0 | 0 | 0 | 3 | 3 |
| 4 | 0 | 0 | 1 | 1 | 0 | 2 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | 1 | 0 | 2 |
| 1 | 4 | 5 | 2 | 3 | 4 | 18 |
| N | 4 | 6 | 3 | 5 | 7 | 25* |
| M | 1.0 | 1.2 | 2.0 | 1.8 | 2.7 | |

Chi-square = 17.1; P = greater than .10

TABLE XXIII

FREQUENCY DISTRIBUTION OF PRACTICE NO. 3: "ESTABLISHES RAPPORT"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 0 | 1 | 0 | 11 | 12 |
| 4 | 1 | 0 | 6 | 1 | 5 | 13 |
| 3 | 0 | 2 | 7 | 2 | 4 | 15 |
| 2 | 0 | 5 | 3 | 2 | 2 | 12 |
| 1 | 12 | 6 | 5 | 0 | 7 | 30 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.2 | 1.7 | 2.8 | 2.8 | 3.4 | |

Chi-square = 53.5; P = less than .01

TABLE XXIV

FREQUENCY DISTRIBUTION OF PRACTICE NO. 4: "WHOLE CLASS IN RECALL"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 2 | 0 | 0 | 10 | 12 |
| 4 | 4 | 1 | 5 | 1 | 6 | 17 |
| 3 | 0 | 1 | 5 | 1 | 3 | 11 |
| 2 | 0 | 4 | 5 | 1 | 0 | 10 |
| 1 | 9 | 5 | 7 | 1 | 10 | 32 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.9 | 2.3 | 2.4 | 2.6 | 3.2 | |

Chi-square = 36.5; P = less than .01

TABLE XXV

FREQUENCY DISTRIBUTION OF PRACTICE NO. 10: "USES MENTAL HYGIENE VALUES"

| | P | F | G | VG | E | N |
|---|---|---|---|---|---|---|
| 5 | 0 | 0 | 1 | 1 | 8 | 10 |
| 4 | 0 | 1 | 2 | 1 | 6 | 10 |
| 3 | 0 | 0 | 3 | 0 | 2 | 5 |
| 2 | 0 | 1 | 3 | 2 | 0 | 6 |
| 1 | 13 | 11 | 13 | 1 | 13 | 51 |
| N | 13 | 13 | 22 | 5 | 29 | 82 |
| M | 1.0 | 1.3 | 1.9 | 2.8 | 2.9 | |

Chi-square = 36.1; P = less than .01

* Practice on new form only.

**For accurate value of P based upon coarser groupings, see Appendix E in original thesis Johns Hopkins University.

frequency distribution for Practice No. 34 "Provides for elastic endings." It is true that as many as 14 of the 19 teachers rated 5 by the observer were the principals' E teachers, but too many of the G, VG, and E teachers (27 out of the 48 teachers rated 1) were rated 1 by the observer. Because these 1 ratings were weighted lightly in the M row and because the concentration of 17 E teachers in ratings 5 and 4 weights heavily, Figure 6 shows a fairly steep slope and a rise of 2.1 units. The chi-square figure of 30.3 is fairly high with a probability that there are two* chances in 100 that this figure could have been reached by chance. Thus, ratings on the practice of providing elastic endings are fairly well related to our principals' concept of effective discipline.

Practice No. 9 "Uses subject matter appeal as a control" shows a higher relationship in all three analyses. The frequency distribution in Table VII has 30 of the 33 teachers rated 5 in the G, VG, and E groups. The same ability teachers contain 10 of the 12 cases rated 4. On the lower end of the distribution, 15 of the 21 teachers rated 1 are P or F calibre. Naturally, this type of distribution, when weighted and averaged, yields a steep slope as shown by Figure 7. The chi-square, although not as high as in Practice No. 5 above, is high enough (36.4) to give a probability figure that there are only 7 chances in 1000 that a chi-square as large as this could have occurred by chance. Thus, we can safely say that the observer's rating on the practice of using subject matter appeal as a control is very definitely associated with the principals' ratings of effective discipline.

Practice No. 16 "Meets disciplinary situation," which would seem to be closely related to the principals' concept of discipline, does not give as clear a picture. Perhaps this comes about because of the limited number of cases available; perhaps not. At any rate, the frequency distribution in Table VIII has six of the 15 E teachers rated 1 on this practice, with the other nine spread throughout the other observer ratings. Speaking for the relationship, it is true that 17 of the 26 teachers rated 1 are either P or F teachers, and three of the five teachers rated 5 are E teachers. Figure 8 bears out the sparse relationship by exhibiting a line that tends definitely toward the horizontal in slope. The spread of the 15 E teachers throughout all 5 observer ratings with the preponderance at the bottom accounts for the mean for this group being lower than the mean for the G group, for the horizontal slope of the line at the end, and for a rise of only 1.3 units altogether. The chi-square is also only fair, with a test figure of 25.6 and a probability indicating that there are as many as 6* chances in 100 that this figure could have been reached by chance. All in all, a probability of the size suggests a relationship of questionable significance. In this connection, it should be noted that the reliability figure for this trait is only .46.

The next practice, Practice No. 12 "Avoids lulls or idleness," shows a better relationship. The frequency distribution in Table IX reveals that 19 of the 29 E teachers are found in ratings 4 and 4, and 11 of the 13 P teachers are rated 1. The relationship isn't as close as it might be because there are 6 E teachers rated 1, but the general picture is good. This is borne out by the graph in Figure 9 which shows a definite vertical tendency in the slope of the line, rising two whole units from P to E. The chi-square is also fairly good (31.2), high enough to show only 1-1/2* chances in 100 that this figure could have occurred by chance. The observer's rating of practice of avoiding lulls or idleness, then, seems fairly definitely associated with the principals' idea of effective discipline.

Practice No. 11 "Plans worthwhile work to keep each pupil challenged constantly" shows even a closer relationship between our two variables. The frequency distribution in Table X reveals that 17 of the 25 teachers rated 5 were of either VG or E calibre, and 16 of the 25 teachers rated 1 were either P or F teachers. The graph in Figure 10 bears out this high relationship with a steep slope, rising 2.5 units from P through E. For further corroboration, one need but consult the chi-square figure of 41.4, high enough to warrant a probability that there are only 7 chances in 10,000 that this relationship could have occurred by chance. All told, the practice of planning worthwhile work so that each pupil is challenged constantly seems very definitely associated with the principals' concept of effective discipline.

The next practice, Practice No. 13 "Classroom manner free rather than restricted," takes us downhill again in the matter of relationship. Although it is true that 6 of the 9 teachers rated 5 are E teachers, the rest of the 29 E teachers are distributed throughout the other 4 observer ratings, with 7 of them coming at the bottom with a rating of 1. The graph in Figure 11 shows this tapering off at the high end and accounts for the lack of steepness in the 1.4 units of rise involved in the graph. The chi-square completes the poor picture with a low of 24.4 and a probability indicating that there are as many as 8** chances in 100 that this figure could have occur-

* Accurate value of P is less than .01.

** Accurate value P is between .05 and .01, but because of low objectivity this practice was not included in the final discussion of the eight most significant practices.

red by chance. Thus, the practice of having a free classroom manner seems very poorly related to the principals' estimate of discipline. Limited relation is to be expected with low objectivity (. 51) and reliability (. 57).

Practice No. 22 "Explains points carefully" seems just a bit better although the limited amount of cases makes for poor prognosis. Table XII shows 4 of the 5 teachers rated 5 are E teachers, but 7 E teachers hit the 1 row. The small number of cases make for a misleading graph, for the slope in Figure 12 is perhaps steeper than it would be with additional data, and the 1. 9 units of rise are not very significant considering the small numbers involved. The chi-square of 28. 8 is only mediocre with a probability that there are 2-1/2 chances in 100 that this could have occurred by chance. This is an intermediate degree of significance, all the more surprising in view of the low objectivity (. 50) and low reliability (. 48).

Practice No. 23 "Answers points of confusion patiently" goes up again slightly in relationship. Table XIII reveals that 13 of the 19 teachers rated 5 are E teachers, and that 21 of the 26 P and F teachers are rated 1. However, the high relationship is damaged in this same row for we find 10 G teachers and 11 E teachers are also rated 1. The slope of the line in the graph in Figure 13 shows a fair amount of steepness and encompasses 1. 7 units of rise. Likewise, the chi-square of 31. 7 does allow for the probability that there are 1-1/2 * chances in 100 that this figure could have been reached by chance. We can safely say that there is some relationship between the observer's ratings and the principals' ratings where the practice of answering points of confusion patiently is concerned. This holds in spite of the low objectivity (. 44) and low reliability (. 49).

Continuing this practice-by-practice triple analysis of the relationship between the observer's ratings on the practice for the teachers concerned and the principals' ratings of the disciplinary abilities of these teachers, we come next to Practice No. 6 "Cares for physical conditions of the room," which shows a closer over-all picture of relationship than the last three practices discussed. Table XIV indicates that 19 of the 29 E teachers reached 5 or 4 ratings, 10 of the 13 P teachers were rated 1, and again the 22 G teachers were distributed throughout all five observer ratings to furnish the middle trend. This picture is illustrated in Figure 14 where the 2. 2 units of rise stretch out in a fairly steep slope, except for the irregular 7 cases out of the 13 F teachers who were rated 5 or 4. The reader will notice that this irregularity accounts for the high start of the broken line of the smoothed data and detracts from the steepness of the graph as a whole. The chi-square is high (40. 7) with a probability low enough to indicate that there are only 7 chances in 10, 000 that this figure could have been reached by chance. Ratings on the practice of taking care of the physical conditions of the room seem , then, to be definitely associated with the principals' concept of effective discipline. Again this relation is achieved in the face of low objectivity (. 42) and reliability (. 54).

The next practice, Practice No. 2 "Begins when willingness is shown," shows the closest relationship to the disciplinary ratings of the teachers by the principals of all 24 practices analyzed. Table XV reveals that 14 of the 19 teachers rated 5 are E teachers, as many as 19 of the 21 teachers rated 4 are eigher G, VG, and E teachers, and 11 of the 13 P teachers find their commensurate ratings in row 1. To indicate the close relationship even more, 9 of the 13 F teachers are found with either 1 or 2 ratings. The graph in Figure 15 shows an almost straight-line relationship rising 2. 4 units. The chi-square of 55. 5 is the highest figure obtained accompanied by the lowest probability (1 chance in 10,000 that this figure could have come about by chance! ). It is evident that whatever elements go to make up the practice of beginning instruction only when pupil willingness is shown, are much the same elements that form the basis for principals' judgment of teachers' disciplinary ability; that whatever it was the observer saw in teacher technique or teacher personality that made a teacher wait for a work attitude, or for irrelevant noise to subside, seem to be the very things that make principals rate teachers high in disciplinary ability. Certainly we can say that the practice of beginning when willingness is shown is one practice closely associated with our principals' concept of effective discipline. In this connection, it should be noted that although this trait had fair reliability (. 62), it had low objectivity (. 40).

From the apex of relationship reached in the previous practice, we go severely downhill in a consideration of the next practice, No. 31 "Shows awareness of differences in ability." Table XVI does show that 9 of the 13 teachers rated 5 are E teachers, and 12 of the 13 P teachers do have a 1 rating. However, the rest of the distribution breaks the relationship down: 16 of the 22 G teachers, 4 of the 5 VG teachers, and 14 of the 29 E teachers are rated 1! This horizontal trend is amply illustrated in Figure 16 where the 1. 4 units of rise show the lack of steepness even

* Accurate value of P is less than .01 but because of low objectivity this practice was not included in the final discussion of the eight most significant practices

when the data is smoothed to counteract the undue effect of the 5 VG teachers. The chi-square hits the compromise figure of 30.2 with a probability indicating that there are as many as 2* chances in 100 that this figure could have been reached by chance. Like Practice No. 2, this trait has fair reliability (.78) but low objectivity (.37).

A similar picture is obtained with Practice No. 21 "Provides for individual instruction" as evidenced in Table XVII. Here again, although 7 of the 9 teachers rated 5 are E teachers, and 12 of the 13 P teachers find themselves in the 1 row, 15 of the 22 G teachers and 16 of the 29 E teachers also find themselves in the 1 row. The graph in Figure 17 illustrated this leveling off with the 1.1 units of rise making for definite lack of steepness. The chi-square is higher than might be expected (37.8) with 4 chances in 1000 that this figure could have been reached by chance. All in all, providing for individual instruction is another practice only moderately associated with the principals' concept of effective discipline. This trait again has fair reliability (.68) but low objectivity (.37).

The next practice, Practice No. 24 "Makes pupils feel the importance of the work they are engaged in," gives a mixed picture. Table XVIII suggests considerable relationship. There are 15 E teachers in the 5 or 4 rows, and 22 of the 26 P or F teachers are rated 1. Yet, here again we have 9 of the 22 G teachers and 9 of the 29 E teachers in the 1 row. The 1.7 units of rise mount fairly steeply in Figure 18, especially when viewed with the broken line of the smoothed data. However, the chi-square takes a dip to the low figure of 24.6 with a probability that there are as many as 8** chances in 100 that this relationship could have been reached by pure chance. Such a figure suggests very questionable significance. This lack of relationship should be considered in the light of the fairly high reliability (.68) and low objectivity (.34).

Practice No. 17 "Puts progress of the whole group above individual considerations" shows little relation. We see in Table XIX that 9 of the 12 E teachers involved in this distribution are rated 3 or below, 2 of the 4 VG teachers are rated 1, and 5 of the 10 G teachers are rated either 2 or 1. Indicating some relationship, however, is the fact that all 6 P teachers are rated 1 and account for the good low start of the line of relationship in Figure 19. The reader will notice that after 1.0 unit of rise is jumped from the P to the F point, the rest of the o.6 units or rise flatten out. The chi-square bears out the lack of relationship with the low figure of 19.0 and a probability indicating almost 27 chances in 100 that this is purely a chance relationship. The lack of relationship could be due to the limited number of cases (38), to the very limited objectivity (.32) or to the limited reliability (.60). On the basis of the evidence available the practice of putting progress of the whole group above individual considerations seem insignificantly related to our principals' ratings of effective discipline.

Practice No. 1 "Makes presence felt in a pleasant way" gives a much better picture of relationship. Table XX reveals that 17 of the 29 E teachers achieve a rating of 5 or 4, and 13 of the 22 G teachers center around the middle of the observer's ratings. However, the distribution is given a sharp jolt by having 5 of the 13 P teachers land in the 4 row. This accounts for the high start of the graph in Figure 20 and for the fact that there are only 1.3 units of rise making for a less steep slope than would obtain if the P teachers were all rated low by the observer. The chi-square reaches a fairly high figure of 37.2 with a probability that there are no more than 6 chances in 1000 that the observer's ratings in this practice are related to the principals' ratings by chance alone. Thus, the practice of entering a room in a pleasant manner and setting the emotional tone of the period by making one's presence felt pleasantly seems fairly related to the principals' concept of effective discipline. This degree of relationship may seem surprising in view of the low objectivity (.32) and reliability (.51).

Practice No. 15 "Displays sense of humor for relevant fun" hits the nadir or relationship. Table XXI reveals that only 4 of the 29 E teachers are rated 5 in this practice. Furthermore, 15 of these E teachers evidently display no sense of humor at all and are rated 1. In addition, 15 of the 22 G teachers and 3 of the 5 VG teachers display no sense of humor. The only cases that save Figure 21 from an almost perfect horizontal line are the 12 E teachers rated 3 or better, making the mean for E column 2.3 and allowing for 0.8 units of rise. The chi-square of 15.9 is the lowest obtained in all the data and shows the good probability that the relationship is a chance

* Accurate value of P is between .10 and .05.

**Accurate value of P is less than .01 but because of the low objectivity this practice was not included in the final discussion of the eight most significant practices.

one (46* chances in 100). This lack of relationship could be due to the low objectivity (.37) and reliability (.38). This apparent lack of relationship could also represent the real situation which might appear even with more reliable and objective observation.

A practice showing almost as poor relationship is No. 35 "Gives summary or has pupils give it," but perhaps this is due to the limited number of cases available for study in this practice. At any rate, Table XXII reveals that 4 of the 7 E teachers rated on the practice achieve a 1 rating, 3 of the 5 VG teachers do the same thing, and 2 of the 3 G teachers achieve likewise. Indicating some relationship is the fact that 9 of the 10 P and F teachers are also rated 1 and give the graph in Figure 22 some slope and 1.7 units of rise. The chi-square is very low (17.1) with the good probability that the relationship is a chance one (40 chances in 100), at least on the basis of the meager data available. Such a low relationship is to be expected with a reliability of only .18.

Practice No. 3 "Takes steps to establish rapport" takes us to the extreme opposite end of the picture into a world of meaningful relationships almost as Practice No. 2, Table XXIII, exhibits phenomenal correspondence between observer and principal ratings. Of the 12 teachers rated 5, 11 are E teachers; the G teachers center around the middle; and 12 of the 13 P teachers are rated 1. This almost perfect picture is spoiled only by the fact that 7 of the 29 E teachers are also rated 1. The graph in Figure 23 illustrates the closeness of the relationship by the steepness of the slope and the fact that there are 2.2 units of rise involved. A chi-square of 53.5 indicates the slight probability of a chance relationship (1-1/2 in 10,000). Whatever teachers do in approaching a closer bond with their pupils, whether it be a stressing of common purposes or a skillful adjustment of inter-personal relations, there is something in or about this establishing of rapport that seems definitely related to the principals' estimates of disciplinary ability. The practice may be poorly worded or too intangible to be precisely defined[14] (as seems to be indicated by lack of objectivity .073, noted in the previous section), but establishing rapport is definitely associated with the principals' concept of effective discipline.

Practice No. 4 "Has whole class active in recall" is much less closely related. Table XXIV reveals that although 10 of the 29 E teachers do achieve a rating of 5 in the practice, another 10 of the E teachers are rated 1; although 9 of the 13 P teachers are rated 1, the other 4 achieve a rating of 4. This last accounts for the high start of the graph in Figure 24, with a middle leveling off because of the 12 G teachers (out of 22) rated either 2 or 1. It is only because 6 more E teachers achieve a 4 rating to help out the 10 in the 5 row that slope rises at all at the end, finishing the distribution of the 1.3 units rise concerned in the whole graph. The chi-square of 36.5 is fairly high, indicating only 7 chances in 1000 for a chance relationship. All told, the observer's rating of the practice of having the whole class active in recall seems only moderately related to the principals' ratings of effective discipline. This relation appears in spite of the extremely low objectivity (.07) and the low reliability (.50).

The last practice open for analysis, Practice No. 10 "Uses mental hygiene values for each pupil," shows a slightly better all-round picture of relationship even though the chi-square is about the same. Table XXV indicates that the difference comes in the P column where all 13 teachers are found to be rated 1. This lower end is helped further by the fact that 12 of the 13 F teachers are rated either 2 or 1. The upper end of the distribution does have 14 of the 29 E teachers rated either 5 or 4, but also has 13 of the remainder rated 1, along with 13 of the 22 G teachers also rated 1. The graph in Figure 25 gives a steep slope, largely because of the low start and the fact that there are 1.9 units of rise to cover. The chi-square of 36.1 indicates a fair relationship, with only 7-1/2 chances in 1000 that it could have occurred by chance. The observer's rating of the practice of using mental hygiene considerations in the treatment of pupils in interpersonal relations seems fairly well associated with the principals' concept of effective discipline. It should be noted that this trait although fairly reliable (.67) is almost completely lacking in objectivity (.05).

The triple analysis of the closeness of the association between the observer's ratings on the various practices and the principals' ratings of the disciplinary abilities of the 82 teachers of the study, presented in this section, show the following practices to be most closely related to these principals' concept of effective discipline: making use of equipment and visual aids, develops classroom mechanics, uses subject matter appeal as a control, plans worthwhile work to keep each pupil challenged constantly, takes care of

* Accurate value of P is between .05 and .01 but because of low objectivity this practice was not included in the final discussion of the eight most significant practices.

14. See suggestions for interpretation at end of this section for other possible explanations.

physical conditions of the room, begins when willingness is shown, and takes steps to establish rapport. The question of how objective and reliable are those same practices is best answered by the synthesis performed in the interpretation that follows and in the next section.

Suggestions for Interpretation

The following considerations may be helpful in interpreting the data in the summary table of the next section.

In the case of a practice coming out with low reliability, low objectivity, and a low chi-square, it may be ascribed to the possibility that the lack of objectivity made it impossible to achieve any reliability, or the possibility that there was no relationship at all.

In the case of low reliability, low objectivity, and a high chi-square, it is possible that there was something the author did or saw on first observation which agrees with the principals' rating, but not something which can be seen every time or by every observer. Neither the author nor any other observers could be sure of what to look for.

In the case of low reliability, high objectivity, and low chi-square, the possible explanations are the lack of reliability or the possibility of no relationship at all.

In the case of low reliability, high objectivity, and high chi-square, one possible explanation is that the author and the supervisors were rating the same thing, this same thing agreeing with the principals' rating but not susceptible to observation.

In the case of high reliability, low objectivity, and low chi-square, the explanation may be simply that here is the something the author sees that he saw consistently but that nobody else saw, and that had no relation to discipline.

In the case of high reliability, low objectivity, and high chi-square, the possible explanation is that what the author sees consistently agrees with the principals' ratings but that the description of this practice is not adequate for other observers.

In the case of high reliability, high objectivity, and low chi-square, the explanation is that what the author sees other observers see, but that there is no relation to the principals' ratings.

Finally, in the case of high reliability, high objectivity, and high chi-square, we may have the desideratum: what the author sees other observers see and the practice is closely related to the principals' ratings in discipline.

To synthesize these data, the results on all three counts have been put into a summary table (see Table XXVI).[15] A study of this table reveals three practices outstanding in objectivity, reliability, and relationship: Practice No. 5 "Makes use of equipment and visual aids" (.88, .68, 44.5),* Practice No. 7 "Develops classroom mechanics" (.68, .54, 34.2),* and Practice No. 9 "Uses subject matter appeal as a control" (.59, .66, 36.4).* Some of the other practices have higher correlations in either reliability or objectivity, or higher chi-squares, but fall down in one aspect or another. There are, however, three additional practices that are fairly significant in all three aspects, although not as significant as the first three mentioned above. These three are marked off with the first three from the other practices purely arbitrarily, with an objectivity r of .59 as the lowest mark-off point acceptable: Practice No. 12 "Avoids lulls or idleness (.72, .48, 31.2),* Practice No. 18 "Uses simple devices, a look or a smile" (.86, .67, 26.7),* and Practice No. 34 "Provides for elastic endings" (.73, .52, 30.3).*

In order to discover some possible relationships among the most significant practices, it was at first thought that a factor analysis might be used. However, a spot-check analysis of the total correlations for eight traits (two additional traits, No.'s 16 and 20 were included for this purpose), arranged in descending order according to the Spearman idea of factor analysis, yielded such highly irregular end coefficients that any thought of a thorough factor analysis was abandoned. In fact, as it turned out, so few traits met all the criteria that the factor analysis was not pursued beyond a mere table of inter-correlations. These appear in Table XXVII.

The correlations range all the way from .34 to .84, and, as was mentioned above, there seems to be no apparent single factor accounting for the relationship. However, an examination of the practices themselves and their descriptions indicate that all eight are, by function, techniques of classroom teaching which, the author maintains, are practiced by the skillful teacher.

Viewing the study as a whole, one may summarize the results as follows:

1. Making use of equipment and visual aids, developing classroom mechanics, and using subject matter appeal as a control are

15. Since the charts and the graphs show much the same trend as indicated by the chi-squares, the discussion of relationship is generalized around the chi-square data alone.

* For these and recalculated values of chi-square, see Appendix E in original thesis Johns Hopkins University.

TABLE XXVI

SUMMARY TABLE OF OBJECTIVITY, RELIABILITY, AND CORRECTED PROBABILITY FOR EACH PRACTICE

| Practice | A Objectivity | B Reliability | C Corrected P |
|---|---|---|---|
| 1. Makes presence pleasant | .32 | .51 | Less than .01 |
| 2. Begins when willingness is shown | .40 | .62 | Less than .01 |
| 3. Takes steps to establish rapport | .07 | .76 | Less than .01 |
| 4. Has whole class active in recall | .07 | .50 | Less than .01 |
| 5. Uses equipment and visual aids | .88 | .68 | Less than .01 |
| 6. Physical conditions of room | .42 | .54 | Less than .01 |
| 7. Develops classroom mechanics | .68 | .54 | Less than .01 |
| * | | | |
| 9. Subject matter appeal as control | .59 | .66 | Less than .01 |
| 10. Uses mental hygiene values | .05 | .67 | Less than .01 |
| 11. Plans worthwhile work | .54 | .43 | Less than .01 |
| 12. Avoids lulls or idleness | .72 | .48 | Less than .01 |
| 13. Classroom manner free | .51 | .57 | Between .05 and .01 |
| * | | | |
| 15. Displays sense of humor | .37 | .38 | Between .05 and .01 |
| 16. Meets disciplinary situation | .90 | .46 | Less than .01 |
| 17. Whole group above individual | .32 | .60 | Greater than .10 |
| 18. Uses simple devices | .86 | .67 | Less than .01 |
| * | | | |
| 20. Uses blackboard well | .86 | .76 | Less than .01 |
| 21. Provides for individual instruction | .37 | .68 | Less than .01 |
| 22. Explains points carefully | .49 | .50 | ** |
| 23. Answers confusion patiently | .44 | .49 | Less than .01 |
| 24. Makes pupils feel importance of work | .34 | .68 | Less than .01 |
| * | | | |
| 31. Aware of ability differences | .37 | .78 | Between .10 and .05 |
| * | | | |
| 34. Provides for elastic endings | .73 | .52 | Less than .01 |
| 35. Gives summary | .71 | .18 | ** |

* Other practices deleted because of insufficient data; either the practice was included too late or it was not possible to be present in many observations.

**Recalculation of chi-square impossible because of lack of cases.

TABLE XXVII

TABLE OF INTER-CORRELATIONS AMONG THE EIGHT MOST SIGNIFICANT PRACTICES

| | Develops classroom mechanics | Uses subject matter appeal | Avoids lulls or idleness | Meets disciplinary situation | Uses simple devices | Uses blackboard | Provides for elastic endings |
|---|---|---|---|---|---|---|---|
| Makes use of equipment and visual aids | .60 | .54 | .52 | .51 | .56 | .62 | .41 |
| Develops classroom mechanics | | .60 | .61 | .56 | .61 | .53 | .34 |
| Uses subject matter appeal as a control | .60 | | .84 | .69 | .52 | .61 | .52 |
| Avoids lulls or idleness | .61 | .84 | | .70 | .63 | .62 | .50 |
| Meets disciplinary situation | .55 | .69 | .70 | | .67 | .73 | .63 |
| Uses simple devices | .61 | .52 | .63 | .67 | | .58 | .46 |
| Uses blackboard | .53 | .61 | .62 | .73 | .58 | | .40 |
| Provides for elastic endings | .34 | .52 | .50 | .63 | .46 | .40 | |

three objective, reliable practices closely associated with the principals' ratings of effective discipline.

2. Avoiding lulls or idleness, using simple devices, providing elastic endings, are three addtional practices that are reliable and objective, and moderately associated with the principals' concept of effective discipline.

3. Beginning when willingness is shown and taking steps to establish rapport have elements in them very closely associated with the principals' concept of effective discipline, but are not objective enough to be considered significant as they are now worded or discerned, or as traits that exist.

4. The other practices in this study are either too unreliable, too subjective, or too poorly related to the principals' concept of effective discipline to be considered significant for the purposes of this study.

## Summary and Conclusions

### A. Summary

The purpose of this investigation was to study descriptively the dozen or so vital practices that make for a good disciplinarian in a teacher, those practices that are closely associated with principals' concepts of effective discipline.

A list of thirty-eight practices thought to be possibly important were drawn up into a rating chart used in the observation of the teachers concerned in the study. Eighty-two teachers from the Baltimore secondary schools (three junior and three senior) voluntarily participated and were observed once; sixty-six of these were observed a second time; and two observed a third time to make a total of one hundred fifty observations. Twenty addtional teachers were observed once in the objectivity check by the observer and one of four supervisors. The academic ability of the classes and the disciplinary ability of the teachers observed were fairly evenly distributed throughout. The eighty-two teachers were rated in discipline control by their respective principals after the observations were made by the observer, and these principals' ratings were matched with the observer's ratings on each practice for each teacher to give the data on relationship explored in the last part of the last section. The principals' ratings were expressed in terms of E (Excellent), VG (Very Good), G (Good), F (Fair), P (Poor), and the observer's ratings were expressed in terms of 5, 4, 3, 2, and 1 to denote the descending order of the performance of any practice.

The results were handled in three different ways. First, the objectivity of each practice was determined by the results of the joint observations of the twenty addtional teachers by the observer and either one of the four supervisors used for this purpose. Then, the reliability of each practice was determined from the observations of the sixty-six teachers seen twice. Finally, the principals' ratings of the teachers were plotted against their ratings by the observer on each practice to give rise to a series of frequency distributions, a series of straight-line graphs, and a series of chi-squares and their probabilities, all of which should indicate the closeness of relationship between the two variables.

### B. Conclusions

On the basis of the data revealed in this study, and on the basis of an analysis of those data, one may conclude that:

1. The practice of using all available equipment and visual aids to embellish and enrich a lesson so as to interest and promote the learning growth of pupils is closely associated with effective discipline.

2. The practice of routinizing various classroom procedures such as the distribution of paper, copying of the assignment, entering or leaving the room, etc., in developing the mechanics of the classroom as well-formed habits is another practice closely associated with effective discipline.

3. The practice of presenting the subject matter in a vital and enthusiastic manner, of making the subject matter appealing to the point that the interest and the enthusiasm catches on enough so that the subject matter itself acts as a check or control to incipient misbehavior is also closely associated with effective discipline.

4. A number of other practices describing skillful techniques of classroom teaching, such as keeping the class moving without idleness, using simple devices, and providing elastic endings, are moderately associated with effective discipline.

### C. Subjects for Further Study

There are a number of interestingly provocative aspects of this study that should form fruitful subjects in further study:

1. An attempt should be made to discover the elements in such practices as "begins when willingness is shown" and "takes steps to establish rapport" which account for the very close association between them and the principals' ratings in effective discipline.

2. The more intengible practices such as "a free classroom manner" and "a sense of humor" need intense study.

3. Finally, there is the legion of questions which continually reappear in or around a study of this type and which need answering: Are there essential differences in disciplinary practice followed between all boy and all girl classes, between before gymnasium and after gymnasium classes, between before lunch and after lunch classes, between first period and last period classes, between bright and dull classes, between rural and city classes, between special classes and regular classes, between private school classes and public school classes, and between classes taught by veteran teachers and those classes taught by beginners?

*A New*

# TEACHER PERSONNEL RESEARCH

**MONOGRAPH NO. 3**

by

DAVID G. RYANS
LAWRENCE P. BLUM
KENNETH R. DOANE
LOIS G. NEMEC

**This new monograph (132 pages) contains four reports in which answers are sought to such questions as:**

- **How well do teachers perform on the National Teachers Examination?**
- **How do persons preparing to teach compare with those of other selected professions?**
- **What is current practice in the education of teachers in the United States?**
- **How is teacher education related to teacher certification?**

***Ready Now*** ***Price $2.50***

## Dembar Publications, Inc.

**Madison 3, Wisconsin**

*A New*

# Educational Statistics Primer

**Written by students for students**

- Statistics self-taught for the non-mathematical student
- Who desires a more intelligent reading of technical papers in education, and
- Who desires to make easy statistical calculations with speed and accuracy

This pocket edition of statistics explains and illustrates calculation procedures for:

Measures of central position

Measures of variability

Zero order correlation and simple regression

Errors of measurement and sampling

Multiple correlation and regressions

Ready now . . . . . Price $1.25

**DEMBAR PUBLICATIONS, Inc.**

**Madison 3, Wisconsin**